

Optimal reinsurance under risk and uncertainty

ALEJANDRO BALBÁS,^{*}BEATRIZ BALBÁS,[†]RAQUEL BALBÁS[‡]AND ANTONIO HERAS[§]

Abstract. This paper deals with the optimal reinsurance problem if both insurer and reinsurer are facing risk and uncertainty, though the classical uncertainty free case is also included. The insurer and reinsurer degrees of uncertainty do not have to be identical. The decision variable is not the retained (or ceded) risk, but its sensitivity with respect to the total claims. Thus, if one imposes strictly positive lower bounds for this variable, the reinsurer moral hazard is totally eliminated.

Three main contributions seem to be reached. Firstly, necessary and sufficient optimality conditions are given. Secondly, the optimal contract is often a bang-bang solution, *i.e.*, the sensitivity between the retained risk and the total claims saturates the imposed constraints. For some special cases the optimal contract might not be bang-bang, but there is always a bang-bang contract as close as desired to the optimal one. Thirdly, the optimal reinsurance problem is equivalent to other linear programming problem, despite the fact that risk, uncertainty, and many premium principles are not linear. This may be important because linear problems are easy to solve in practice, since there are very efficient algorithms.

KEY WORDS. Risk and Uncertainty, Moral Hazard, Optimal Reinsurance and Optimality Conditions, Bang-Bang Solution, The Optimal Reinsurance Linear Problem.

A.M.S. CLASSIFICATION SUBJECT. 91B30, 90C48.

JEL CLASSIFICATION. G22.

1. INTRODUCTION

Since Borch (1960) and Arrow (1963) published their celebrated seminal papers, the optimal reinsurance problem has been addressed by many authors and under many different risk measurement methods and premium principles. Recent approaches are, amongst many others, Kaluszka (2005), Cai and Tan (2007) and Chi and Tan (2013).

^{*}University Carlos III of Madrid. CL. Madrid 126. 28903 Getafe (Madrid, Spain). alejandro.balbas@uc3m.es

[†]University of Castilla la Mancha. Avda. Real Fábrica de Seda, s/n. 45600 Talavera (Toledo, Spain). beatriz.balbas@uclm.es

[‡]University Complutense of Madrid. Department of Actuarial and Financial Economics. Somosaguas-Campus. 28223 Pozuelo de Alarcón (Madrid, Spain). raquel.balbas@ccee.ucm.es

[§]University Complutense of Madrid. Department of Actuarial and Financial Economics. Somosaguas-Campus. 28223 Pozuelo de Alarcón (Madrid, Spain). aheras@ccee.ucm.es

Usually, researchers consider the insurer point of view, though the reinsurer viewpoint may be also incorporated (Cai *et al.*, 2012, Cui *et al.*, 2013, etc.). An interesting survey about the State of the Art in 2009 may be found in Centeno and Simoes (2009).

All the papers above assume that the statistical distribution of claims is known. Nevertheless, measurement errors or lack of complete information may provoke discrepancies between the real and the estimated probabilities of the states of nature, generating uncertain (also called ambiguous) frameworks. Actuarial and financial literature is recently paying significant attention to those cases where the probabilities of the scenarios are not totally known. Interesting examples are, among many others, portfolio management (Zhu and M. Fukushima, 2009), equilibrium in asset markets (Bossaerts *et al.*, 2010) and optimal stopping (Riedel, 2009).

The first objective of this paper is to incorporate ambiguity in the optimal reinsurance problem, though many results will be also new in the uncertainty free setting. Both insurer and reinsurer may be ambiguous, but their degrees of ambiguity do not have to be identical. Since the reinsurer information about the reinsured set of policies could be lower than the information of the insurer, it seems natural to assume that the reinsurer ambiguity is higher, but we will not impose this hypothesis because we will not need it. According to the empirical evidence and the famous Ellsberg paradox, agents usually reflect ambiguity aversion, so we will accept this assumption in our analysis. Though there are other recent approaches (Maccheroni *et al.*, 2006), the worst-case principle properly incorporates the ambiguity aversion (Gilboa and Schmeidler, 1989), and therefore our analysis will deal with this principle when considering the insurer expected wealth, the insurer global risk (integrating uncertainty too) and the reinsurer premium principle. Actually, all of the papers above deal with ambiguity by means of a worst-case approach.

Stop-loss or closely related contracts frequently solve the optimal reinsurance problem. These solutions have been often criticized by both theoretical researchers and practitioners. In practice, reinsurers will rarely accept these solutions due to the lack of incentives of the insurer to verify claims beyond some thresholds. Our second objective will be to overcome this caveat. Consequently, the insurer decision variable will be the (almost everywhere) mathematical derivative of the retained risk with respect to the global claims, rather than the retained risk itself. With this modification we can impose positive lower bounds to this decision variable, and therefore contracts reflecting spreads with null derivative (flat behavior of the retained risk with respect to the global claims) become unfeasible. In other words, the usual reinsurer moral hazard is eliminated with this approach.

The paper is organized as follows. Section 2 will present the general framework, the set of priors, the properties of the insurer risk measure (integrating uncertainty), the properties of the reinsurance premium principle (which may incorporate the rein-

surer uncertainty) and the general optimal reinsurance problem we are going to deal with. We will point out how our approach contains most of the usual cases and extends them all if ambiguity arises. Section 3 will be devoted to dealing with two dual approaches. Theorems 4 and 5 will provide us with two alternative dual problems, as well as two different families of necessary and sufficient optimality conditions. It is worth to point out that one of the duals is linear.

The optimality conditions will generate two different ways permitting us to linearize the optimal reinsurance problem. The first one is the introduction of a linear optimization problem generated by a dual solution. This method will allow us to prove Theorem 7 in Section 4, which will show that the optimal contract is in the closure of a set composed of convex combinations of bang-bang contracts, *i.e.*, contracts such that the derivative of the retained risk with respect to the total one saturates the imposed constraints. A clear consequence is that in many classical approaches one must find stop-loss or closely related optimal contracts. In our less restrictive framework the optimal retention will be often a bang-bang solution.

Section 5 explores a second linearization procedure. In order to simplify the exposition the focus is on the Robust Conditional Value at Risk (robust $CVaR$) as an insurer risk/ambiguity measure and a reinsurer instrument to generate the premium principle. The method applies for much more situations, but selecting one important case we significantly shorten the paper. Furthermore, the $CVaR$ is becoming more and more popular among researchers and practitioners due to its interesting properties (Ogryczak and Ruszczyński, 2002).

Since one of the two duals of Section 3 is linear, we will construct the double-dual (dual of the dual) optimal reinsurance problem in Section 5, which is linear. We will prove that the solution of the double-dual will directly lead to the optimal reinsurance contract. This seems to be a very important property because there are many efficient algorithms solving linear problems in both finite-dimensional and infinite-dimensional frameworks (Anderson and Nash, 1987). Besides, linear problems often lead to extreme solutions, which explains why the non linear optimal reinsurance problem may be solved by a bang-bang retention.

The last section of the paper summarizes the most important conclusions, emphasizing the two main novelties (uncertainty introduction and moral hazard elimination) and the three main contributions (necessary and sufficient optimality conditions, bang-bang solutions and double-dual linear problems).

Throughout the paper we will need several mathematical notions about topological spaces, Banach and Hilbert spaces, strong and weak convergences, etc. Some of them will be briefly summarized, but further discussions may be found in Luenberger (1969), Kelly (1975), Rudin (1973) and (1987), or Anderson and Nash (1987), amongst others.

2. THE OPTIMAL REINSURANCE PROBLEM

Let (Ω, \mathcal{F}) be a measurable space, Ω denoting the set of states of nature at a future date T . Assume that the insurer is ambiguous (or reflects uncertainty) with respect to the probabilities of the states, so we cannot select a unique probability measure on \mathcal{F} . Let the \mathcal{F} -measurable function (random variable) $y_0 : \Omega \rightarrow \mathbb{R}$ represent the total cost (claims) that the insurer will pay within the period $[0, T]$. It is obvious that y_0 cannot achieve negative values. Since the existence of an upper bound M for y_0 is obvious (claims cannot be higher than the value of the insured goods) y_0 is $[0, M]$ -valued. Moreover, every probability measure on \mathcal{F} consistent with the insurer ambiguity will generate a probability measure on the Borel σ -algebra \mathcal{B} of $[0, M]$. Thus, the role of y_0 may be played by the identity function $I : [0, M] \rightarrow [0, M]$, and the uncertainty level will be given by a set \mathcal{P}_U^0 of probability measures (or set of priors) on \mathcal{B} .

2.1. The set of priors. Next, let us give the main properties that \mathcal{P}_U^0 will have to satisfy. Denote by \mathcal{P} the set of probability measures on \mathcal{B} and fix $\mathbb{P}_0 \in \mathcal{P}$. Consider the Hilbert space $L^2(\mathbb{P}_0)$, which is composed of those random variables x whose square has finite expectation with respect to \mathbb{P}_0 and which is endowed with the norm

$$\|x\|_{(2, \mathbb{P}_0)} = \left[\int_0^M x^2(t) d\mathbb{P}_0(t) \right]^{1/2}, \quad x \in L^2(\mathbb{P}_0). \quad (1)$$

Similarly, consider the usual Lebesgue measure on $[0, M]$ and the classical Hilbert space $L^2[0, M]$, whose well-known norm is given by

$$\|x\|_2 = \left[\int_0^M x^2(t) d(t) \right]^{1/2}, \quad x \in L^2[0, M]. \quad (2)$$

We will assume the existence of $R \in L^2(\mathbb{P}_0)$, $R \geq 1$, such that

$$\mathcal{P}_U^0 = \left\{ p \in \mathcal{P}; 0 \leq \frac{dp}{d\mathbb{P}_0} \leq R \right\}, \quad (3)$$

$\frac{dp}{d\mathbb{P}_0}$ denoting the Radon-Nikodym derivative of p with respect to \mathbb{P}_0 . If $\mathbb{E}_p(\cdot)$ denotes mathematical expectation with respect to a probability measure p , and, in particular, $\mathbb{E}_{\mathbb{P}_0}(\cdot)$ denotes mathematical expectation with respect to \mathbb{P}_0 , the set of priors (3) may be also given by

$$\mathcal{P}_U = \left\{ f \in L^2(\mathbb{P}_0); 0 \leq f \leq R, \mathbb{E}_{\mathbb{P}_0}(f) = 1 \right\}, \quad (4)$$

since we can obviously identify every \mathbb{P}_0 -continuous probability measure $p \in \mathcal{P}_U^0$ with its Radon-Nikodym derivative $f = \frac{dp}{d\mathbb{P}_0}$. In other words, the insurer uncertainty will

be identified with a subset of the interval $[0, R] \subset L^2(\mathbb{P}_0)$. Though it is an abuse of language, we will also say the (4) is the set of the insurer priors.

The set of priors (3) is very general. Firstly, if $R = 1$ then the non-ambiguous (or uncertainty free) case will be included in our framework, because $\mathcal{P}_U^0 = \{\mathbb{P}_0\}$ (or $\mathcal{P}_U = \{1\}$) will become a singleton, and the probabilities of the states of nature will be known. Secondly, if $R \neq 1$, then \mathbb{P}_0 may be interpreted as a estimated probability measure containing possible errors, which makes the insurer incorporate the “spread” $R \geq 1$ indicating the estimation accuracy (the accuracy increases as $R - 1 \geq 0$ decreases). Thirdly, for some $t \in [0, M]$ and some $p \in \mathcal{P}_U^0$ we have that $t \notin \text{Sp}(p)$ may hold, $\text{Sp}(p)$ denoting the support of p . Hence, t would not be possible if p were the real probability measure. In other words, we allow the insurer to be ambiguous with respect to both the probabilities of the states of nature and the set of states of nature too.

Notice that $\mathcal{P}_U \subset [0, R]$ is convex, closed and bounded, which implies that \mathcal{P}_U is a convex and weakly-compact subset of $L^2(\mathbb{P}_0)$ (Alaoglu’s Theorem and Hahn-Banach Theorem). Actually, the most important results of this paper would remain true if we considered a more general framework and allowed \mathcal{P}_U to be an arbitrary convex and weakly-compact subset of $L^2(\mathbb{P}_0)$, but the choice of (4) simplifies the mathematical exposition.

2.2. The decision variable and the problem constraints. Next let us introduce an auxiliary functional which will be useful in order to guarantee that the insurer and reinsurer risks are comonotonic. Consider $J : L^2[0, M] \rightarrow L^2(\mathbb{P}_0)$ given by

$$J(x)(t) = \int_0^t x(s) d(s), \quad x \in L^2[0, M], \quad t \in [0, M]. \quad (5)$$

It is easy to see that $J(x)$ is a real-valued continuous function on $[0, M]$ for every $x \in L^2[0, M]$, so $J(x)$ is bounded and therefore $J(x) \in L^2(\mathbb{P}_0)$. Actually, one can also find many classical references showing that, out of Lebesgue null sets, x is the first order derivative of $J(x)$.

Proposition 1. *The functional $J : L^2[0, M] \rightarrow L^2(\mathbb{P}_0)$ is linear and continuous.*

Proof. J is trivially linear, so let us see its continuity. We only have to prove the existence of $k > 0$ with

$$\|J(x)\|_{(2, \mathbb{P}_0)} \leq k \|x\|_2 \quad (6)$$

for every $x \in L^2[0, M]$. Obviously,

$$|J(x)(t)| \leq \int_0^t |x(s)| d(s) \leq \int_0^M |x(s)| d(s) = \|x\|_1$$

for every $x \in L^2[0, M]$ and every $t \in [0, M]$, $\|\cdot\|_1$ representing the usual norm of the Banach space $L^1[0, M] \supset L^2[0, M]$. Hence,

$$\int_0^M J(x)(t)^2 d\mathbb{P}_0(t) \leq \|x\|_1^2 \int_0^M d\mathbb{P}_0(t) = \|x\|_1^2,$$

which, according to (1), leads to

$$\|J(x)\|_{(2, \mathbb{P}_0)} \leq \|x\|_1 \quad (7)$$

for every $x \in L^2[0, M]$. Since the natural inclusion of $L^2[0, M]$ in $L^1[0, M]$ is continuous, there exists $k > 0$ such that

$$\|x\|_1 \leq k \|x\|_2 \quad (8)$$

for every $x \in L^2[0, M]$. Thus, (6) trivially follows from (7) and (8). \square

An important role will be also played by the linear and continuous functional $J^* : L^2(\mathbb{P}_0) \rightarrow L^2[0, M]$, adjoint of J . If

$$\begin{aligned} \prec x, x' \succ &= \int_0^M x(t) x'(t) dt \\ \text{and} \\ \mathbb{E}_{\mathbb{P}_0}(x''x''') &= \int_0^M x''(t) x'''(t) d\mathbb{P}_0(t), \end{aligned} \quad (9)$$

for $x, x' \in L^2[0, M]$ and $x'', x''' \in L^2(\mathbb{P}_0)$ represent the usual bilinear products in $L^2[0, M]$ and $L^2(\mathbb{P}_0)$ respectively, it is known that J^* is characterized by

$$\mathbb{E}_{\mathbb{P}_0}(J(x)z) = \prec x, J^*(z) \succ \quad (10)$$

for every $x \in L^2[0, M]$ and every $z \in L^2(\mathbb{P}_0)$. Thus,

$$\begin{aligned} \mathbb{E}_{\mathbb{P}_0}(J(x)z) &= \int_0^M J(x)(t) z(t) d\mathbb{P}_0(t) = \int_0^M \left(\int_0^t x(s) ds \right) z(t) d\mathbb{P}_0(t) \\ &= \int_0^M x(s) \left(\int_s^M z(t) d\mathbb{P}_0(t) \right) d(s) = \int_0^M x(s) J^*(z)(s) d(s) = \prec x, J^*(z) \succ, \end{aligned}$$

which leads to

$$J^*(z)(s) = \int_s^M z(t) d\mathbb{P}_0(t), \quad z \in L^2(\mathbb{P}_0), \quad s \in [0, M]. \quad (11)$$

The identity function $I : [0, M] \rightarrow [0, M]$ trivially satisfies $I = J(1)$, 1 denoting the obvious constant function on $[0, M]$. As said above, I represents the total claims within the time interval $[0, T]$. Suppose that a reinsurance contract divides these claims according to the amounts $y \in L^2(\mathbb{P}_0)$ (retained risk) and $I - y \in L^2(\mathbb{P}_0)$

(ceded risk). There are two usual constraints in the literature. Firstly, $0 \leq y \leq I$ and $0 \leq I - y \leq I$ must hold. Secondly, I , y and $I - y$ must be comonotonic, since otherwise the insurer/reinsurer could be facing moral hazard. In order to guarantee that both restrictions hold, we will represent the reinsurance contract with the variable $x \in L^2[0, M]$. Once the insurer selects x , the retained risk will be $y = J(x) \in L^2(\mathbb{P}_0)$, while the ceded risk will be $I - y = I - J(x) = J(1 - x) \in L^2(\mathbb{P}_0)$ (see (5)). If $0 \leq x \leq 1$ (and therefore $0 \leq 1 - x \leq 1$) then the retained risk $J(x)$ and the ceded risk $J(1 - x)$ will trivially satisfy the two required constraints. Indeed, out of Lebesgue null sets, x is the first order derivative (sensitivity) of $J(x)$ with respect to the total claims, and $1 - x$ is the derivative of $J(1 - x)$. Thus, I , $J(x)$, and $J(1 - x)$ are obviously comonotonic because they are non decreasing functions (their first order derivatives are non negative). $x = 0$ leads to the full reinsurance (the whole risk is ceded) and $x = 1$ leads to the null reinsurance (the whole risk is retained).

Ideas above allows us to address a classical caveat often pointed out by several authors. A stop-loss contract, or a closely related one, frequently solves the optimal reinsurance problem. Thus, the reinsurer faces moral hazard once global claims become higher than a contract threshold. More generally, the reinsurer could consider that $x(t) = 0$ in some interval $[a, b] \subset [0, M]$ might provoke moral hazard again, since the retained risk $y = J(x)$ will remain constant in $[a, b]$. This drawback may be overcome by imposing the constraint $H \leq x$, where $0 \leq H \leq 1$ is a measurable function representing the lowest sensitivity (first order derivative) between the retained risk and the total claims that the reinsurer can accept. We could also accept $H = 1$ in some $[a, b] \subset [0, M]$, which means that the reinsurer will not accept to pay claims within the spread $[a, b]$.

Bearing in mind all the ideas above, the feasible set of the optimal reinsurance problem will be

$$F = \{x \in L^2[0, M]; H \leq x \leq 1\},$$

x representing sensitivity between retained risk and claims, and $H \in L^2[0, M]$, $0 \leq H \leq 1$ representing the lowest acceptable sensitivity.

2.3. The risk function. Let us follow the approach of Artzner *et al.* (1999) and Rockafellar *et al.* (2006) in order to introduce the notion of risk function. Thus, throughout this paper a risk measure will be a function $\rho : L^2(\mathbb{P}_0) \rightarrow \mathbb{R}$ such that there exists a convex and weakly-compact subset Δ_ρ of $L^2(\mathbb{P}_0)$ with

$$\rho(y) = \text{Max} \{-\mathbb{E}_{\mathbb{P}_0}(yz); z \in \Delta_\rho\} \quad (12)$$

for every $y \in L^2(\mathbb{P}_0)$. If (12) holds then it may be proved that Δ_ρ is unique, and $z \in \Delta_\rho$ if and only if

$$-\mathbb{E}_{\mathbb{P}_0}(yz) \leq \rho(y)$$

for every $y \in L^2(\mathbb{P}_0)$. Furthermore, (12) holds if and only if ρ is continuous, sub-additive ($\rho(y_1 + y_2) \leq \rho(y_1) + \rho(y_2)$ for $y_i \in L^2(\mathbb{P}_0)$, $i = 1, 2$) and homogeneous ($\rho(\lambda y) = \lambda \rho(y)$ for $\lambda \geq 0$ and $y \in L^2(\mathbb{P}_0)$). We will also impose ρ to be \tilde{E}_ρ -translation invariant for some $\tilde{E}_\rho \in \mathbb{R}$ ($\rho(y + k) = \rho(y) - \tilde{E}_\rho k$ if $k \in \mathbb{R}$ and $y \in L^2(\mathbb{P}_0)$). It is equivalent to the fulfillment of

$$\mathbb{E}_{\mathbf{P}_0}(z) = \tilde{E}_\rho \quad (13)$$

for every $z \in \Delta_\rho$. Summarizing, we have:

Definition 1. Let $\tilde{E}_\rho \in \mathbb{R}$ be a real number. $\rho : L^2(\mathbb{P}_0) \rightarrow \mathbb{R}$ is said to be a \tilde{E}_ρ -translation invariant risk measure if the two equivalent conditions below hold;

a) $\rho : L^2(\mathbb{P}_0) \rightarrow \mathbb{R}$ is norm-continuous, \tilde{E}_ρ -translation invariant, sub-additive and homogeneous.

b) There exists a convex and weakly-compact set $\Delta_\rho \subset L^2(\mathbb{P}_0)$ such that (12) holds for every $y \in L^2(\mathbb{P}_0)$ and (13) holds for every $z \in \Delta_\rho$. \square

We will not prove the equivalence between Conditions a) and b) above because analogous proofs may be found in several papers (for instance, see Balbás *et al.*, 2013). Notice that (12) implies that ρ is weakly lower semi-continuous.

There are many risk measures in practice. Actually, every expectation bounded risk measure and every deviation measure (Rockafellar *et al.*, 2006) satisfy Definition 1 above with $\tilde{E}_\rho = 1$ and $\tilde{E}_\rho = 0$, respectively. The coherent risk measures of Artzner *et al.* (1999) also satisfy Definition 1 with $\tilde{E}_\rho = 1$. As said in the introduction, the Ellsberg paradox indicates the presence of ambiguity aversion, which may be incorporated by dealing with the worst-case principle (Gilboa and Schmeidler, 1989). A consistent way to consider this principle is to define “the robust extension of a risk measure relative to the set of priors \mathcal{P}_U^0 ”. For instance, the robust *CVaR* with confidence level $0 \leq \mu < 1$ relative to the set of priors \mathcal{P}_U^0 will be given by

$$RCVaR_{(\mathcal{P}_U^0, \mu)}(y) := \text{Max} \{ CVaR_{(p, \mu)}(y) ; p \in \mathcal{P}_U^0 \} \quad (14)$$

for every $y \in L^2(\mathbb{P}_0)$, $CVaR_{(p, \mu)}(y)$ denoting the usual *CVaR* of y if p is the selected probability measure and μ is the level of confidence. Theorem 2 below shows that this definition is consistent.

Theorem 2. Suppose that $0 \leq \mu < 1$. Then, $CVaR_{(p, \mu)}(y)$ exists for every $p \in \mathcal{P}_U^0$ and every $y \in L^2(\mathbb{P}_0)$, the maximum in (14) exists for every $y \in L^2(\mathbb{P}_0)$, and $RCVaR_{(\mathcal{P}_U^0, \mu)}$ satisfies Definition 1b) with $\tilde{E}_{RCVaR_{(\mathcal{P}_U^0, \mu)}} = 1$ and

$$\Delta_{RCVaR_{(\mathcal{P}_U^0, \mu)}} = \left\{ z \in L^2(\mathbb{P}_0) ; \mathbb{E}_{\mathbf{P}_0}(z) = 1 \text{ and } \exists f \in \mathcal{P}_U \text{ with } 0 \leq z \leq \frac{f}{1-\mu} \right\}. \quad (15)$$

In particular, (12) leads to

$$\left\{ \begin{array}{l} RCVaR_{(\mathcal{P}_U^0, \mu)}(y) = \text{Max} - \mathbb{E}_{\mathbf{P}_0}(yz) \\ \mathbb{E}_{\mathbf{P}_0}(f) = 1 \\ \mathbb{E}_{\mathbf{P}_0}(z) = 1 \\ 0 \leq z \leq \frac{f}{1-\mu}, \quad f \leq R \\ z \in L^2(\mathbb{P}_0) \text{ and } f \in L^2(\mathbb{P}_0) \text{ are decision variables} \end{array} \right. \quad (16)$$

for every $y \in L^2(\mathbb{P}_0)$.

Proof. First let us prove that the set $\Delta_{RCVaR_{(\mathcal{P}_U^0, \mu)}}$ given by (15) is convex. Indeed, suppose that z_1 and z_2 belong to this set and take $0 \leq \lambda \leq 1$. If f_1 and f_2 are the obvious measurable functions given by (15), then $z = \lambda z_1 + (1 - \lambda) z_2$ and $f = \lambda f_1 + (1 - \lambda) f_2$ trivially satisfy the conditions guaranteeing that $z \in \Delta_{RCVaR_{(\mathcal{P}_U^0, \mu)}}$.

$\Delta_{RCVaR_{(\mathcal{P}_U^0, \mu)}}$ is norm-bounded. Indeed, (15) implies that $\|z\|_{(2, \mathbf{P}_0)} \leq \left\| \frac{R}{1-\mu} \right\|_{(2, \mathbf{P}_0)}$ for every $z \in \Delta_{RCVaR_{(\mathcal{P}_U^0, \mu)}}$.

$\Delta_{RCVaR_{(\mathcal{P}_U^0, \mu)}}$ closed. Indeed, suppose that the sequence $(z_n)_{n=1}^\infty \subset \Delta_{RCVaR_{(\mathcal{P}_U^0, \mu)}}$ converges to z in $L^2(\mathbb{P}_0)$ and let us prove that $z \in \Delta_{RCVaR_{(\mathcal{P}_U^0, \mu)}}$. Let $(f_n)_{n=1}^\infty \subset L^2(\mathbb{P}_0)$ the obvious sequence. Since $(f_n)_{n=1}^\infty \subset \mathcal{P}_U$ and this set is weakly compact, there exists $f \in \mathcal{P}_U$ which is a weak agglomeration point. Thus (z, f) is a weak agglomeration point of $(z_n, f_n)_{n=1}^\infty$ and therefore $0 \leq z \leq \frac{f}{1-\mu}$, $f \leq R$ trivially follow from $0 \leq z_n \leq \frac{f_n}{1-\mu}$, $f_n \leq R$ for every $n \in \mathbb{N}$.

$\Delta_{RCVaR_{(\mathcal{P}_U^0, \mu)}}$ is weakly compact. Indeed, $\Delta_{RCVaR_{(\mathcal{P}_U^0, \mu)}}$ is weakly closed because it is closed and convex (Hahn-Banach Theorem), so the Alaoglu's Theorem implies that $\Delta_{RCVaR_{(\mathcal{P}_U^0, \mu)}}$ is weakly compact because it is norm-bounded.

$\Delta_{RCVaR_{(\mathcal{P}_U^0, \mu)}}$ being convex and weakly compact, we have that the maximum in the right hand side of (16) obviously exists. Denote by $\rho(y)$ this maximum. It is known (Rockafellar *et al.*, 2006) that for $p \in \mathcal{P}_U^0$

$$CVaR_{(p, \mu)}(y) = \text{Max} \left\{ -\mathbb{E}_p(yz); \mathbb{E}_p(z) = 1, 0 \leq z \leq \frac{1}{1-\mu} \right\} \quad (17)$$

holds whenever y is integrable with respect to p . If $y \in L^2(\mathbb{P}_0)$ then $|y| R$ is integrable with respect to \mathbb{P}_0 and

$$\mathbb{E}_p(|y|) = \int_0^M |y| dp = \int_0^M |y| \frac{dp}{d\mathbb{P}_0} d\mathbb{P}_0 \leq \int_0^M |y| R d\mathbb{P}_0 < \infty$$

for every $p \in \mathcal{P}_U^0$, so y is integrable with respect to p and $CVaR_{(p,\mu)}(y)$ exists.

The theorem will be obviously proved if we show that $\rho(y)$ equals the right hand side of (14). Denote by $RCVaR_{(\mathcal{P}_U^0,\mu)}(y) \leq \infty$ the supremum in the right hand side of (14). If $p \in \mathcal{P}_U^0$ (17) implies the existence of $0 \leq \tilde{z} \leq \frac{1}{1-\mu}$ with

$$1 = \mathbb{E}_{\mathbf{p}}(\tilde{z}) = \mathbb{E}_{\mathbf{P}_0}\left(\tilde{z} \frac{dp}{d\mathbf{P}_0}\right)$$

and

$$CVaR_{(p,\mu)}(y) = -\mathbb{E}_{\mathbf{p}}(y\tilde{z}) = -\mathbb{E}_{\mathbf{P}_0}\left(y\tilde{z} \frac{dp}{d\mathbf{P}_0}\right).$$

$z = \tilde{z} \frac{dp}{d\mathbf{P}_0} \in \Delta_{RCVaR_{(\mathcal{P}_U^0,\mu)}}$ because $f = \frac{dp}{d\mathbf{P}_0}$ satisfies the required conditions. Hence, $CVaR_{(p,\mu)}(y) = -\mathbb{E}_{\mathbf{P}_0}(yz) \leq \rho(y)$ and therefore $RCVaR_{(\mathcal{P}_U^0,\mu)}(y) \leq \rho(y)$.

The opposite inequality also holds. Indeed, if $\rho(y) = -\mathbb{E}_{\mathbf{P}_0}(yz)$ with $z \in \Delta_{RCVaR_{(\mathcal{P}_U^0,\mu)}}$, then take the obvious f and the probability measure $dp = f d\mathbf{P}_0$ satisfies $p \in \mathcal{P}_U^0$. Moreover, $\tilde{z} = \frac{z}{f}$ (take $\frac{0}{0} = 0$) satisfies $0 \leq \tilde{z} \leq \frac{1}{1-\mu}$ and

$$\mathbb{E}_{\mathbf{p}}(\tilde{z}) = \int_0^M \tilde{z} dp = \int_0^M \tilde{z} f d\mathbf{P}_0 = \int_0^M z d\mathbf{P}_0 = \mathbb{E}_{\mathbf{P}_0}(z) = 1,$$

so

$$\begin{aligned} RCVaR_{(\mathcal{P}_U^0,\mu)}(y) &\geq CVaR_{(p,\mu)}(y) \geq -\mathbb{E}_{\mathbf{p}}(y\tilde{z}) = -\int_0^M y\tilde{z} dp = \\ &= -\int_0^M y\tilde{z} f d\mathbf{P}_0 = -\int_0^M yz d\mathbf{P}_0 = -\mathbb{E}_{\mathbf{P}_0}(yz) = \rho(y). \end{aligned}$$

□

Remark 1. The constant random variable $z = 1$ belongs to $\Delta_{RCVaR_{(\mathcal{P}_U^0,\mu)}}$ because $f = 1$ satisfies the conditions of (15). Hence, (16) shows that $RCVaR_{(\mathcal{P}_U^0,\mu)}(y) \geq -\mathbb{E}_{\mathbf{P}_0}(y)$ for every $y \in L^2(\mathbf{P}_0)$, i.e., $RCVaR_{(\mathcal{P}_U^0,\mu)}$ is expectation bounded in the sense of Rockafellar et al. (2006). □

Remark 2. $z \in \Delta_{CVaR_{(\mathbf{P}_0,\mu)}} \implies z \in \Delta_{RCVaR_{(\mathcal{P}_U^0,\mu)}}$ because $f = 1$ satisfies the required conditions. Moreover, If $R \geq 1$ is constant then

$$0 \leq z \leq \frac{f}{1-\mu}, \quad f \leq R \implies 0 \leq z \leq \frac{R}{1-\mu} = \frac{1}{1-\tilde{\mu}}$$

with $\tilde{\mu} = 1 - (\frac{1-\mu}{R})$. Therefore, (16) and (17) trivially lead to

$$CVaR_{(\mathbf{P}_0,\mu)}(y) \leq RCVaR_{(\mathcal{P}_U^0,\mu)}(y) \leq CVaR_{(\mathbf{P}_0,\tilde{\mu})}(y).$$

It is easy to see that $0 \leq 1 - \frac{1}{R} \leq \tilde{\mu} < 1$ and $\tilde{\mu} \geq \mu$. In other words, if $R \geq 1$ is constant then the robust CVaR lays within the spread of two CVaRs. \square

Remark 3. Similarly, one can define a robust extension relative \mathcal{P}_U^0 of many coherent or expectation bounded risk measures, such as the Wang measure, the power transform, the weighted CVaR, etc. All of them will satisfy Definition 1 with $\tilde{E}_\rho = 1$. Robust deviations satisfying $\tilde{E}_\rho = 0$ may be also defined. For instance, the robust absolute deviation will be given by

$$RAD_{\mathcal{P}_U^0}(y) := \text{Max} \left\{ \mathbb{E}_p(|y - \mathbb{E}_p(y)|); p \in \mathcal{P}_U^0 \right\},$$

and the robust standard deviation will be

$$RSD_{\mathcal{P}_U^0}(y) := \text{Max} \left\{ \sqrt{\mathbb{E}_p((y - \mathbb{E}_p(y))^2)}; p \in \mathcal{P}_U^0 \right\}.$$

We will not analyze the deviations above because they can be studied with the methods of Theorem 2. \square

2.4. The premium principle. If $x \in L^2[0, M]$ is the retained risk and $1 - x \in L^2[0, M]$ is the ceded one, then reinsurance price of will be given by

$$\text{Reinsurance_Price} = \Upsilon(J(1 - x)), \quad (18)$$

$\Upsilon : L^2(\mathbb{P}_0) \rightarrow \mathbb{R}$ being a \tilde{E}_Υ -translation invariant risk measure for some $\tilde{E}_\Upsilon \in \mathbb{R}$ (Definition 1). Notice that (18) contains the usual premium principles. Indeed, for a linear principle

$$\Upsilon(J(1 - x)) = \mathbb{E}_{\mathbf{P}_0}(J(1 - x)c), \quad (19)$$

with $c \in L^2(\mathbb{P}_0)$ and $c \geq 0$, it is obvious that Υ satisfies Definition 1b) with $\Delta_\Upsilon = \{-c\}$ and $\tilde{E}_\Upsilon = -\mathbb{E}_{\mathbf{P}_0}(c)$. In particular, the Expected Value Premium Principle (EVPP)

$$\text{Reinsurance_Price} = (1 + \alpha) \mathbb{E}_{\mathbf{P}_0}(J(1 - x)),$$

$\alpha \geq 0$ being the loading rate, is a linear principle with $c = 1 + \alpha$, and therefore it is included in (18). As said above, every deviation measure and every expectation bounded or coherent risk measure ρ satisfies Definition 1, and therefore so does

$$\Upsilon(J(1 - x)) = \alpha \mathbb{E}_{\mathbf{P}_0}(J(1 - x)) + \beta \rho(-J(1 - x)), \quad (20)$$

$\alpha, \beta \geq 0$, with

$$\Delta_\Upsilon = \{-\alpha - \beta z; z \in \Delta_\rho\} \quad (21)$$

and $\tilde{E}_\Upsilon = -\alpha - \beta\tilde{E}_\rho$ (see (13)). Consequently, many classical non-linear premium principles are included in (18) (Wang Principle, Standard Deviation Premium Principle, etc.). More importantly, (18) also includes premium principles involving the reinsurer ambiguity since, for instance, robust extensions of the *CVaR*, the weighted *CVaR*, the Wang measure or the standard deviation may play the role of ρ in (20). Moreover, Υ may be related to the ambiguity level of the reinsurer, which does not necessary equal the insurer ambiguity level. For instance, it may make sense to assume that the reinsurer ambiguity level is higher, and therefore it is characterized by the set of priors $\left\{f \in L^2(\mathbb{P}_0); 0 \leq f \leq \tilde{R}, \mathbb{E}_{\mathbb{P}_0}(f) = 1\right\}$ with $\tilde{R} \geq R$ (see (4)).

2.5. The global problem. The presence of ambiguity aversion and the worst-case principle (Gilboa and Schmeidler, 1989) may also justify the introduction of “the robust expectation relative to the set of priors \mathcal{P}_U ” below

$$\mathbb{E}_{\mathcal{P}_U}(W) := \text{Min} \left\{ \mathbb{E}_p(W); p \in \mathcal{P}_U^0 \right\} = \text{Min} \left\{ \mathbb{E}_{\mathbb{P}_0}(Wf); f \in \mathcal{P}_U \right\} \quad (22)$$

where $W \in L^2(\mathbb{P}_0)$ is the (uncertain) insurer wealth at T , \mathcal{P}_U is the set of feasible priors and $\mathbb{E}_p(W) = \mathbb{E}_{\mathbb{P}_0}(Wf)$ is the expectation of W under $dp = f d\mathbb{P}_0$. Notice that the minimum in (22) exists because

$$L^2(\mathbb{P}_0) \ni f \rightarrow \mathbb{E}_{\mathbb{P}_0}(Wf) \in \mathbb{R}$$

is a weakly continuous function for every $W \in L^2(\mathbb{P}_0)$ and \mathcal{P}_U is a weakly compact set (Weierstrass Theorem). Obviously, (22) leads to

$$-\mathbb{E}_{\mathcal{P}_U}(W) = \text{Max} \left\{ -\mathbb{E}_p(W); p \in \mathcal{P} \right\} = \text{Max} \left\{ -\mathbb{E}_{\mathbb{P}_0}(Wf); f \in \mathcal{P}_U \right\}, \quad (23)$$

for every $W \in L^2(\mathbb{P}_0)$. In other words, $-\mathbb{E}_{\mathcal{P}_U}$ also satisfies Definition 1 with $\tilde{E}_{-\mathbb{E}_{\mathcal{P}_U}} = 1$ and $\Delta_{-\mathbb{E}_{\mathcal{P}_U}} = \mathcal{P}_U$.

Suppose that A denotes the income generated by the sold insurance policies. Bearing in mind the reinsurance contract and (18), the insurer wealth at T will be

$$A - J(x) - \Upsilon(J(1-x)). \quad (24)$$

The insurer may select a risk measure ρ (Definition 1) in order to control the risk of (24). The optimal reinsurance problem will maximize the robust expected wealth $\mathbb{E}_{\mathcal{P}_U}$ (which is equivalent to the minimization of $-\mathbb{E}_{\mathcal{P}_U}$) and will minimize ρ . Thus, bearing in mind the properties of $-\mathbb{E}_{\mathcal{P}_U}$ and ρ (Definition 1), the optimization problem becomes

$$\begin{cases} \text{Min } \rho(A - J(x) - \Upsilon(J(1-x))) = \rho(-J(x)) + \tilde{E}_\rho \Upsilon(J(1-x)) - \tilde{E}_\rho A \\ \text{Min } -\mathbb{E}_{\mathcal{P}_U}(A - J(x) - \Upsilon(J(1-x))) = -\mathbb{E}_{\mathcal{P}_U}(-J(x)) + \Upsilon(J(1-x)) - A \\ H \leq x \leq 1 \end{cases}$$

$x \in L^2[0, M]$ being the decision variable. Since \tilde{E}_ρ and A are constant we can remove $\tilde{E}_\rho A$ and A in the vector optimization problem above without altering the Pareto solutions. Moreover, since (12) and (23) show that $-\mathbb{E}_\mathcal{P}$ and ρ are convex functions on $L^2(\mathbb{P}_0)$, every proper Pareto optimum may be obtained by minimizing a scalar objective

$$\rho(-J(x)) + \tilde{E}_\rho \Upsilon(J(1-x)) + w(-\mathbb{E}_{\mathcal{P}_U}(-J(x)) + \Upsilon(J(1-x))),$$

$w > 0$ being a “weight” for $-\mathbb{E}_{\mathcal{P}_U}$ (Nakayama *et al.*, 1985). Actually, w indicates the relative importance of the robust expected return with respect to the risk (the higher the value of w the higher the importance of $\mathbb{E}_{\mathcal{P}_U}$ for the insurer). Manipulating, the objective function becomes

$$\rho(-J(x)) + w(-\mathbb{E}_\mathcal{P}(-J(x))) + (\tilde{E}_\rho + w) \Upsilon(J(1-x))$$

and the optimal reinsurance problem will be

$$\begin{cases} \text{Min } \rho(-J(x)) + w(-\mathbb{E}_{\mathcal{P}_U}(-J(x))) + (\tilde{E}_\rho + w) \Upsilon(J(1-x)) \\ H \leq x \leq 1 \end{cases}, \quad (25)$$

$x \in L^2[0, M]$ being the decision variable.

3. DUAL APPROACH

Let us introduce an instrumental “Mean Value Theorem” whose proof is omitted because a similar one may be found in Balbás *et al.* (2009).

Lemma 3. *Suppose that Y is a Banach space, Z is its dual, and $K \subset Z$ is a convex and $\sigma(Z, Y)$ –compact set. Suppose that ν is an inner regular probability measure on the Borel σ –algebra of K endowed with the $\sigma(Z, Y)$ –topology. Then, there exists a unique $k_\nu \in K$ such that*

$$\int_K \langle y, z \rangle d\nu(z) = \langle y, k_\nu \rangle$$

holds for every $y \in Y$. □

Let us introduce two dual problems for (25), as well as necessary and sufficient Karush-Kuhn-Tucker like optimality conditions. Standard duality theory does not guarantee the absence of duality gap if infinite-dimensional problems are involved. Duality gaps may occur even in linear programming (Anderson and Nash, 1987). In order to overcome this caveat, first of all we will give a concave dual satisfying the

Slater qualification (Luenberger, 1969) which guarantees the absence of duality gap. Then, we will give a linear dual not satisfying the Slater qualification. Nevertheless, the absence of duality gap with the concave dual will allow us to prove the absence of duality gap with the linear one too.

Theorem 4. *Consider Problem*

$$\left\{ \begin{array}{l} \text{Max} \left\{ \begin{array}{l} - \left(\tilde{E}_\rho + w \right) \prec 1, J^*(\gamma) \succ \\ + \prec H, J^* \left(z + wf + \left(\tilde{E}_\rho + w \right) \gamma \right)^+ \succ \\ - \prec 1, J^* \left(z + wf + \left(\tilde{E}_\rho + w \right) \gamma \right)^- \succ \end{array} \right. \\ f \in \mathcal{P}_U, z \in \Delta_\rho, \gamma \in \Delta_\Upsilon \end{array} \right. , \quad (26)$$

$(f^*, z^*, \gamma^*) \in \mathcal{P}_U \times \Delta_\rho \times \Delta_\Upsilon$ being the decision variable.¹ Then:

a) The minimum of (25) and the maximum of (26) are finite, attainable and identical.

b) Suppose that x^* is (25)-feasible and (f^*, z^*, γ^*) is (26)-feasible. Then, x^* solves (25) and (f^*, z^*, γ^*) solves (26) if and only if, out of Lebesgue null sets of $[0, M]$,

$$\left\{ \begin{array}{l} \mathbb{E}_{\mathbf{P}_0} (J(x^*) f^*) \geq \mathbb{E}_{\mathbf{P}_0} (J(x^*) f), \quad \forall f \in \mathcal{P}_U \\ \mathbb{E}_{\mathbf{P}_0} (J(x^*) z^*) \geq \mathbb{E}_{\mathbf{P}_0} (J(x^*) z), \quad \forall z \in \Delta_\rho \\ \mathbb{E}_{\mathbf{P}_0} (J(1 - x^*) \gamma^*) \leq \mathbb{E}_{\mathbf{P}_0} (J(1 - x^*) \gamma), \quad \forall \gamma \in \Delta_\Upsilon \\ x^*(t) = H(t), \quad J^* \left(z^* + wf^* + \left(\tilde{E}_\rho + w \right) \gamma^* \right) (t) > 0 \\ x^*(t) = 1, \quad J^* \left(z^* + wf^* + \left(\tilde{E}_\rho + w \right) \gamma^* \right) (t) < 0 \end{array} \right. \quad (27)$$

Proof. Let us prove that (25) is solvable. Indeed the feasible set $H \leq x \leq 1$ is weakly-compact because $L^2[0, M]$ is a Hilbert space and therefore reflexive (Alaoglu's Theorem), and the objective function of (25) is weakly lower semi-continuous because it is a composition of the $\sigma(L^2[0, M], L^2[0, M]) - \sigma(L^2(\mathbb{IP}_0), L^2(\mathbb{IP}_0))$ -continuous function J and the $\sigma(L^2(\mathbb{IP}_0), L^2(\mathbb{IP}_0))$ -lower semi-continuous functions ρ , $-\mathbb{E}_{\mathcal{P}}$ and Υ . Therefore, (25) attains an optimal value (Weierstrass' Theorem).

¹As usual, $\phi^+(t) = \text{Max}_t \{\phi(t), 0\}$ and $\phi^-(t) = \text{Max}_t \{-\phi(t), 0\}$ for very $\phi \in L^2[0, M]$ and every $0 \leq t \leq M$. It is easy to see that $\phi = \phi^+ - \phi^-$ and $\phi^+, \phi^- \in L^2[0, M]$. Similar notations may be used in similar situations.

Expressions (12) (which also applies for Υ under the obvious modifications) and (23) show that Problem (25) is equivalent to Problem

$$\begin{cases} \text{Min } \theta_1 + w\theta_2 + (\tilde{E}_\rho + w) \theta_3 \\ \theta_1 \geq \mathbb{E}_{\mathbf{P}_0}(J(x)z), & \forall z \in \Delta_\rho \\ \theta_2 \geq \mathbb{E}_{\mathbf{P}_0}(J(x)f), & \forall f \in \mathcal{P}_U \\ \theta_3 \geq \mathbb{E}_{\mathbf{P}_0}(J(x-1)\gamma), & \forall \gamma \in \Delta_\Upsilon \\ H \leq x \leq 1, \theta_1, \theta_2, \theta_3 \in \mathbb{R} \end{cases} \quad (28)$$

$(\theta_1, \theta_2, \theta_3, x) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times L^2[0, M]$ being the decision variable. The first, second and third constraints of (28) are valued in the spaces of continuous functions on the weakly compact spaces $K = \Delta_\rho$, $K = \mathcal{P}_U$ and $K = \Delta_\Upsilon$ respectively. Following Luenberger (1969), the associated dual variables (ν_1, ν_2, ν_3) must be inner regular non-negative measures on the Borel σ -algebras of the compact sets above, and the Lagrangian function is

$$\begin{aligned} \mathcal{L}(\theta_1, \theta_2, \theta_3, x, \nu_1, \nu_2, \nu_3) = & \theta_1 \left(1 - \int_{\Delta_\rho} d\nu_1\right) + \\ & w\theta_2 \left(1 - \int_{\mathcal{P}_U} d\nu_2\right) + (\tilde{E}_\rho + w) \theta_3 \left(1 - \int_{\Delta_\Upsilon} d\nu_3\right) + \\ & \int_{\Delta_\rho} \mathbb{E}_{\mathbf{P}_0}(J(x)z) d\nu_1(z) + w \int_{\mathcal{P}_U} \mathbb{E}_{\mathbf{P}_0}(J(x)f) d\nu_2(f) + \\ & (\tilde{E}_\rho + w) \int_{\Delta_\Upsilon} \mathbb{E}_{\mathbf{P}_0}(J(x)\gamma) d\nu_3(\gamma) - (\tilde{E}_\rho + w) \int_{\Delta_\Upsilon} \mathbb{E}_{\mathbf{P}_0}(J(1)\gamma) d\nu_3(\gamma). \end{aligned} \quad (29)$$

Since (ν_1, ν_2, ν_3) is dual-feasible if and only if

$$\text{Inf } \{\mathcal{L}(\theta_1, \theta_2, \theta_3, x, \nu_1, \nu_2, \nu_3); H \leq x \leq 1, \theta_1, \theta_2, \theta_3 \in \mathbb{R}\} > -\infty$$

(Luenberger, 1969), ν_1 , ν_2 and ν_3 must be probability measures (total mass equal to one), and the dual objective becomes (see (10))

$$\begin{cases} \text{Inf } \begin{cases} \int_{\Delta_\rho} \mathbb{E}_{\mathbf{P}_0}(J(x)z) d\nu_1(z) + w \int_{\mathcal{P}_U} \mathbb{E}_{\mathbf{P}_0}(J(x)f) d\nu_2(f) + \\ (\tilde{E}_\rho + w) \int_{\Delta_\Upsilon} \mathbb{E}_{\mathbf{P}_0}(J(x)\gamma) d\nu_3(\gamma) \\ - (\tilde{E}_\rho + w) \int_{\Delta_\Upsilon} \mathbb{E}_{\mathbf{P}_0}(J(1)\gamma) d\nu_3(\gamma) \end{cases} \\ H \leq x \leq 1 \end{cases}.$$

As in Balbás *et al.* (2009), Lemma 3 with $Y = Z = L^2(\mathbb{P}_0)$ guarantees that the dual solution is achieved in a vector of three Dirac Deltas, so the dual variable may simplify

to $(z, f, \gamma) \in \Delta_\rho \times \mathcal{P}_U \times \Delta_\Upsilon$, and the dual objective becomes (see (9), (10) and (11))

$$\begin{aligned}
\phi(z, f, \gamma) &= -\left(\tilde{E}_\rho + w\right) + \mathbb{E}_{\mathbf{P}_0}(J(1)\gamma) + \\
&\quad \text{Inf}_{H \leq x \leq 1} \left\{ \mathbb{E}_{\mathbf{P}_0}(J(x)z) + w\mathbb{E}_{\mathbf{P}_0}(J(x)f) + \left(\tilde{E}_\rho + w\right) \mathbb{E}_{\mathbf{P}_0}(J(x)\gamma) \right\} \\
&= -\left(\tilde{E}_\rho + w\right) \mathbb{E}_{\mathbf{P}_0}(J(1)\gamma) + \\
&\quad \text{Inf}_{H \leq x \leq 1} \left\{ \prec x, J^*(z) \succ + w \prec x, J^*(f) \succ + \left(\tilde{E}_\rho + w\right) \prec x, J^*(\gamma) \succ \right\} \\
&= -\left(\tilde{E}_\rho + w\right) \prec 1, J^*(\gamma) \succ + \text{Inf}_{H \leq x \leq 1} \left\{ x, J^*\left(z + wf + \left(\tilde{E}_\rho + w\right)\gamma\right) \succ \right\}.
\end{aligned} \tag{30}$$

The dual problem is

$$\begin{cases} \text{Max } \phi(z, f, \gamma) \\ (z, f, \gamma) \in \Delta_\rho \times \mathcal{P}_U \times \Delta_\Upsilon \end{cases} \tag{31}$$

The absence of duality gap between (25) (or (28)) and (31) and the solvability of (31) hold due to the Slater condition (Luenberger, 1969), which is trivially fulfilled by (28).

The infimum of (30) is obviously achieved at

$$x(t) = \begin{cases} H(t), & \text{if } J^*\left(z + wf + \left(\tilde{E}_\rho + w\right)\gamma\right)(t) > 0 \\ H(t) \leq x(t) \leq 1, & \text{if } J^*\left(z + wf + \left(\tilde{E}_\rho + w\right)\gamma\right)(t) = 0 \\ 1, & \text{if } J^*\left(z + wf + \left(\tilde{E}_\rho + w\right)\gamma\right)(t) < 0 \end{cases} \tag{32}$$

and equals

$$\begin{aligned}
&\int_{J^*(z+wf+(\tilde{E}_\rho+w)\gamma)(t)>0} \left(H J^*\left(z + wf + \left(\tilde{E}_\rho + w\right)\gamma\right)(t) \right) dt + \\
&\int_{J^*(z+wf+(\tilde{E}_\rho+w)\gamma)(t)=0} J^*\left(z + wf + \left(\tilde{E}_\rho + w\right)\gamma\right)(t) dt + \\
&\int_{J^*(z+wf+(\tilde{E}_\rho+w)\gamma)(t)<0} 0 dt \\
&= \prec H, \text{Max}_t \left\{ J^*\left(z + wf + \left(\tilde{E}_\rho + w\right)\gamma\right)(t), 0 \right\} \succ \\
&\quad - \prec 1, \text{Max}_t \left\{ -J^*\left(z + wf + \left(\tilde{E}_\rho + w\right)\gamma\right)(t), 0 \right\} \succ.
\end{aligned}$$

Thus, (30) and (31) lead to the dual objective given in (26) and *a*) is proved.

In order to prove b), owing to (32) the necessary and sufficient primal and dual optimality conditions for (28) become (Luenberger, 1969)

$$\left\{ \begin{array}{ll} \theta_1^* = \mathbb{E}_{\mathbf{P}_0} (J(x^*) z^*) \\ \theta_1^* \geq \mathbb{E}_{\mathbf{P}_0} (J(x^*) z), & \forall z \in \Delta_\rho \\ \theta_2^* = \mathbb{E}_{\mathbf{P}_0} (J(x^*) f^*) \\ \theta_2^* \geq \mathbb{E}_{\mathbf{P}_0} (J(x^*) f), & \forall f \in \mathcal{P}_U \\ \theta_3^* = \mathbb{E}_{\mathbf{P}_0} (J(x^* - 1) \gamma^*) \\ \theta_3^* \geq \mathbb{E}_{\mathbf{P}_0} (J(x^* - 1) \gamma), & \forall \gamma \in \Delta_\Upsilon \\ x^*(t) = H(t), & \text{if } J^* \left(z^* + w f^* + \left(\tilde{E}_\rho + w \right) \gamma^* \right) (t) > 0 \\ x^*(t) = 1, & \text{if } J^* \left(z^* + w f^* + \left(\tilde{E}_\rho + w \right) \gamma^* \right) (t) < 0 \end{array} \right. ,$$

and (27) becomes obvious. \square

Remark 4. Despite the fact that (25) is convex but not linear, we will see that it is closely related to several linear problems. Firstly, let us see that (26) may be easily transformed into a linear problem, at least for appropriate choices of \mathcal{P}_U , ρ and Υ . It is important for two reasons. Firstly, there are many algorithms to solve in practice both finite-dimensional and infinite-dimensional linear problems (Anderson and Nash, 1987). Secondly, linear problems often present extreme (or bang-bang) solutions x^* such that $x^*(t) = H(t)$ or $x^*(t) = 1$ much hold out of Lebesgue null sets. This could explain why the stop loss reinsurance, satisfying $x^*(t) = 0$ or $x^*(t) = 1$, frequently solves (25) if $H = 0$. \square

Theorem 5. Consider Problem

$$\left\{ \begin{array}{l} \text{Max} \quad \prec H, \lambda_d \succ - \prec 1, \lambda_u \succ - \left(\tilde{E}_\rho + w \right) \prec 1, J^*(\gamma) \succ \\ J^* \left(z + w f + \left(\tilde{E}_\rho + w \right) \gamma \right) - \lambda_d + \lambda_u = 0 \\ \lambda_d \geq 0, \lambda_u \geq 0 \\ f \in \mathcal{P}_U, z \in \Delta_\rho, \gamma \in \Delta_\Upsilon, \lambda_d \in L^2[0, M], \lambda_u \in L^2[0, M] \end{array} \right. \quad (33)$$

$(f, z, \gamma, \lambda_d, \lambda_u) \in \mathcal{P}_U \times \Delta_\rho \times \Delta_\Upsilon \times L^2[0, M] \times L^2[0, M]$ being the decision variable.

a) The minimum of (25) and the maximum of (33) are finite, attainable and identical.

b) Suppose that x^* is (25)-feasible and $(f^*, z^*, \gamma^*, \lambda_d^*, \lambda_u^*)$ is (33)-feasible. Then, they solve (25) and (33), respectively, if and only if

$$\left\{ \begin{array}{ll} \mathbb{E}_{\mathbf{P}_0} (J(x^*) f^*) \geq \mathbb{E}_{\mathbf{P}_0} (J(x^*) f), & \forall f \in \mathcal{P}_U \\ \mathbb{E}_{\mathbf{P}_0} (J(x^*) z^*) \geq \mathbb{E}_{\mathbf{P}_0} (J(x^*) z), & \forall z \in \Delta_\rho \\ \mathbb{E}_{\mathbf{P}_0} (J(1 - x^*) \gamma^*) \leq \mathbb{E}_{\mathbf{P}_0} (J(1 - x^*) \gamma), & \forall \gamma \in \Delta_\Upsilon \\ \lambda_d^* (x^* - H) = 0 \\ \lambda_u^* (1 - x^*) = 0 \end{array} \right. \quad (34)$$

Proof. As in the proof of Theorem 4, (28) is equivalent to (25). The Lagrangian of (28) may also incorporate the multipliers $\lambda_d, \lambda_u \in L^2[0, M]$, $\lambda_d \geq 0, \lambda_u \geq 0$, related to the constraints $H \leq x \leq 1$. In such a case the Lagrangian becomes

$$\begin{aligned} \mathcal{L}'(\theta_1, \theta_1, \theta_3, x, \nu_1, \nu_2, \nu_3, \lambda_d, \lambda_u) = \\ \mathcal{L}(\theta_1, \theta_1, \theta_3, x, \nu_1, \nu_2, \nu_3) + \prec H - x, \lambda_d \succ + \prec x - 1, \lambda_u \succ, \end{aligned}$$

where \mathcal{L} is given by (29). Besides, the dual variable may be simplified again due to Lemma 3 (see (30)), and the Lagrangian of (28) becomes (see (9), (10) and (11))

$$\begin{aligned} \mathcal{L}'(x, z, f, \gamma, \lambda_d, \lambda_u) &= - \left(\tilde{E}_\rho + w \right) \mathbb{E}_{\mathbf{P}_0}(J(1)\gamma) + \mathbb{E}_{\mathbf{P}_0}(J(x)z) + \\ & w \mathbb{E}_{\mathbf{P}_0}(J(x), f) + \left(\tilde{E}_\rho + w \right) \mathbb{E}_{\mathbf{P}_0}(J(x), \gamma) + \prec H - x, \lambda_d \succ + \prec x - 1, \lambda_u \succ \\ &= - \left(\tilde{E}_\rho + w \right) \prec 1, J^*(\gamma) \succ - \prec 1, \lambda_u \succ + \\ & \prec x, J^*\left(z + wf + \left(\tilde{E}_\rho + w \right) \gamma - \lambda_d + \lambda_u\right) \succ + \prec H, \lambda_d \succ. \end{aligned}$$

The infimum of $\mathcal{L}'(x, z, f, \gamma, \lambda_d, \lambda_u)$ for $x \in L^2[0, M]$ is obviously finite if and only if the first constraint of (33) holds, in which case it equals

$$- \left(\tilde{E}_\rho + w \right) \prec 1, J^*(\gamma) \succ - \prec 1, \lambda_u \succ + \prec H, \lambda_d \succ.$$

Thus, (33) is a dual problem of (28) and (25) (Luenberger, 1969). The Slater qualification does not hold because constraints $H \leq x$ and $x \leq 1$ are valued in $L^2[0, M]$, and the positive cone of this space has void interior. Thus, a duality gap between (25) and (33) might exist, and the supremum of (33) might be strictly lower than the (25)-optimal value. Nevertheless, if we prove that the optimal values of (25) and (33) are reachable and coincide, then the duality gap will vanish, and the solutions of (25) and (33) will be characterized by the complementary slackness conditions (Anderson and Nash, 1987). As in the proof of Theorem 4, the complementary slackness conditions between (28) and (33) coincide with (34), so it only remains to prove that (33) achieves its optimal value and it coincides with the (25)-optimal value. Consider a solution x^* of (25) and a solution (f^*, z^*, γ^*) of (26) (whose existence is guaranteed by Theorem 4a). Consider

$$\lambda_d^* = J^*\left(z^* + wf^* + \left(\tilde{E}_\rho + w\right)\gamma^*\right)^+$$

and

$$\lambda_u^* = J^*\left(z^* + wf^* + \left(\tilde{E}_\rho + w\right)\gamma^*\right)^-.$$

Obviously, $(f^*, z^*, \gamma^*, \lambda_d^*, \lambda_u^*)$ is (33)-feasible, and the dual objective becomes

$$\begin{aligned} & \prec H, \lambda_d^* \succ - \prec 1, \lambda_u^* \succ - \left(\tilde{E}_\rho + w \right) \prec 1, J^*(\gamma^*) \succ \\ & = \prec H, \text{Max} \left\{ J^* \left(z^* + w f^* + \left(\tilde{E}_\rho + w \right) \gamma^* \right), 0 \right\} \succ \\ & - \prec 1, \text{Max} \left\{ -J^* \left(z^* + w f^* + \left(\tilde{E}_\rho + w \right) \gamma^* \right), 0 \right\} + \left(\tilde{E}_\rho + w \right) J^*(\gamma^*) \succ, \end{aligned}$$

which coincides with the (25)-optimal value because there is no duality gap between (25) and (26) (Theorem 4). Hence, since the (25)-optimal value is an upper value for the (33)-objective (Anderson and Nash, 1987), $(f^*, z^*, \gamma^*, \lambda_d^*, \lambda_u^*)$ solves (33) and there is no duality gap between (25) and (33). \square

Remark 5. *Theorem 5 implies that for appropriate \mathcal{P}_U , ρ and Υ the dual solution $(f^*, z^*, \gamma^*, \lambda_d^*, \lambda_u^*)$ of (33) may be obtained in practice by linear programming methods, which is a very good new because interesting algorithms to solve these problems are available in both, finite and infinite dimensions (Anderson and Nash, 1987). Once the dual solution was computed, the optimal retention $J(x^*)$ may be obtained by (34), which is much easier to apply once we know $(f^*, z^*, \gamma^*, \lambda_d^*, \lambda_u^*)$. Besides, if (25) has a linear dual, then this problem should be “almost linear” too. Sections 4 and 5 will show two different methods linearizing Problem (25) (see Problems (36) and (40) below). \square*

4. ON THE EXISTENCE OF BANG-BANG SOLUTIONS

The literature about the optimal reinsurance problem frequently obtains stop-loss or closely related solutions such as stop-loss contracts with an upper bound, given by

$$J(1 - x^*)(t) = (t - a)^+ - (t - b)^+,$$

$0 \leq t \leq M$ and $0 \leq a < b \leq M$.² Obviously, the (Lebesgue almost everywhere) derivative of $J(x^*)$ is given by

$$x^*(t) = \begin{cases} 1, & \text{if } t < a \text{ or } t > b \\ 0, & \text{if } a < t < b \end{cases}$$

and Lemma 6 below will show that this is an extreme point of the (25)-feasible set if $H = 0$. A natural question is how general this result is. Let us prove Theorem 7 below showing that some extreme points are “good approximations” of the (25)-solution for general choices of H and general uncertainty levels for both the insurer and the reinsurer.

²Balbás *et al.* (2009) and (2011) obtain stop-loss optimal contracts even for many problems without comonotonicity restrictions.

Lemma 6. Suppose that x is (25)-feasible. Then, the conditions below are equivalent:

- a) x is a extreme point of the (25)-feasible set $H \leq x \leq 1$.
- b) There exist a Borel-measurable partition $[0, M] = A \cup B$ such that $x(t) = H(t)$ if $t \in A$ and $x(t) = 1$ if $t \in B$.

Henceforth, the feasible solution x above will be called bang-bang retention.

Proof. b) \implies a) is obvious, so let us prove the converse implication. Suppose that a) holds. Consider $C = \{0 \leq t \leq M; H(t) < x(t) < 1\}$ and

$$C_n = \left\{ 0 \leq t \leq M; H(t) + \frac{1}{n} < x(t) < 1 - \frac{1}{n} \right\}$$

for $n \in \mathbb{N}$. Obviously, $C = \cup_{n=1}^{\infty} C_n$. If \mathcal{L} is the Lebesgue measure then $\mathcal{L}(C) = \lim_n \mathcal{L}(C_n)$ because $(C_n)_{n \in \mathbb{N}}$ is an non decreasing sequence. Hence, $\mathcal{L}(C) = 0$ if $\mathcal{L}(C_n) = 0$ for every n . If for some n we have $\mathcal{L}(C_n) > 0$ then $H(t) < x(t) - \frac{1}{n} \leq 1$ and $H(t) \leq x(t) + \frac{1}{n} < 1$ for $t \in C_n$, so functions

$$x_1(t) = x(t) - \frac{1}{n}, \quad x_2(t) = x(t) + \frac{1}{n}$$

for $t \in C_n$ and $x_1(t) = x_2(t) = x(t)$ for $t \notin C_n$ are both (25)-feasible. Thus, $x = \frac{1}{2}x_1 + \frac{1}{2}x_2$ contradicts a). \square

Theorem 7. Suppose that $\tilde{E}_\rho + w \geq 0$. If x^* solves (25) then for every $n \in \mathbb{N}$ there exists a bang-bang feasible retention x_n^* such that the objective function of (25) satisfies

$$\begin{aligned} & \rho(-J(x^*)) + w(-\mathbb{E}_{\mathcal{P}_U}(-J(x^*))) + \left(\tilde{E}_\rho + w\right) \Upsilon(J(1-x^*)) \\ & \leq \rho(-J(x_n^*)) + w(-\mathbb{E}_{\mathcal{P}_U}(-J(x_n^*))) + \left(\tilde{E}_\rho + w\right) \Upsilon(J(1-x_n^*)) + \frac{1}{n}. \end{aligned} \quad (35)$$

Proof. Let (f^*, z^*, γ^*) solve (26). Then, (12), (23), (18) and (27) show that

$$\begin{aligned} \rho(-J(x^*)) &= \mathbb{E}_{\mathbf{P}_0}(J(x^*)z^*) \\ -\mathbb{E}_{\mathcal{P}_U}(-J(x^*)) &= \mathbb{E}_{\mathbf{P}_0}(J(x^*)f^*) \\ \Upsilon(J(1-x^*)) &= -\mathbb{E}_{\mathbf{P}_0}(J(1-x^*)\gamma^*). \end{aligned}$$

The (25)-feasible set is weakly-compact, and therefore, according to the Krein-Milman Theorem, if $n \in \mathbb{N}$ then the linear optimization problem

$$\begin{cases} \text{Min } \mathbb{E}_{\mathbf{P}_0}(J(x)z^*) + w(\mathbb{E}_{\mathbf{P}_0}(J(x)f^*)) - \left(\tilde{E}_\rho + w\right) \mathbb{E}_{\mathbf{P}_0}(J(1-x)\gamma^*) \\ H \leq x \leq 1 \end{cases} \quad (36)$$

satisfies the existence of an extreme point (*i.e.*, a bang-bang retention, see Lemma 6) $H \leq x_n^* \leq 1$ such that

$$\begin{aligned} & \mathbb{E}_{\mathbf{P}_0}(J(x^*)z^*) + w(\mathbb{E}_{\mathbf{P}_0}(J(x^*)f^*)) - \left(\tilde{E}_\rho + w\right) \mathbb{E}_{\mathbf{P}_0}(J(1-x^*)\gamma^*) \leq \\ & \mathbb{E}_{\mathbf{P}_0}(J(x_n^*)z^*) + w(\mathbb{E}_{\mathbf{P}_0}(J(x_n^*)f^*)) - \left(\tilde{E}_\rho + w\right) \mathbb{E}_{\mathbf{P}_0}(J(1-x_n^*)\gamma^*) + \frac{1}{n} \end{aligned}$$

(Anderson and Nash, 1987). Hence, (12), (23) and (18) trivially lead to

$$\begin{aligned} & \mathbb{E}_{\mathbf{P}_0}(J(x_n^*)z^*) + w(\mathbb{E}_{\mathbf{P}_0}(J(x_n^*)f^*)) + \left(\tilde{E}_\rho + w\right) \mathbb{E}_{\mathbf{P}_0}(J(1-x_n^*)(-\gamma^*)) \\ & \leq \rho(-J(x_n^*)) + w(-\mathbb{E}_{\mathcal{P}_U}(-J(x_n^*))) + \left(\tilde{E}_\rho + w\right) \Upsilon(J(1-x_n^*)) \end{aligned}$$

and (35) becomes obvious. \square

Remark 6. Condition $\tilde{E}_\rho + w \geq 0$ will frequently hold. For instance, according to Remark 3, it holds for expectation bounded risk measures, deviation measures, and their robust extensions.

Theorem 7 guarantees that appropriate bang-bang retentions will be as close as desired to the optimal solution. The existence of a bang-bang optimal retention can be proved if the linear problem (36) has an extreme solution. Anderson and Nash (1987) show that this extreme solution can be guaranteed under adequate assumptions. For example, if (25) (or (36)) is a finite-dimensional problem, which holds, for instance, if the total mass of \mathbb{P}_0 is concentrated on a finite subset of $[0, M]$. \square

Remark 7. (Numerical experiments). Conditions (27) of Theorem 4 were useful to prove the existence of bang-bang solutions, and they will be also useful to in order to see that we cannot go beyond Theorem 7 and give closed expressions for sets

$$A = \{t; x^*(t) = H(t)\} \quad \text{and} \quad B = \{t; x^*(t) = 1\},$$

$J(x^*)$ denoting the optimal contract. Under more classical approaches $J(x^*)$ is a stop-loss reinsurance or a closely related one. For a stop-loss contract there exists $M_0 \in [0, M]$ such that

$$A = (M_0, M) \quad \text{and} \quad B = (0, M_0).$$

Nevertheless, in our setting A and B critically depend on the ambiguity set \mathcal{P}_U and the risk functions ρ and Υ , and many different situations may arise. Let us show this fact with simple illustrative examples. Notice that these examples also apply in the non-ambiguous and linear case.

Suppose that $M = 5$, the support of \mathbb{P}_0 is $\{1, 2, 3, 4, 5\}$ and $\mathbb{P}_0(i) = 0.2$, $i = 1, 2, 3, 4, 5$. Suppose that $R = 1$ (non-ambiguous case) and therefore \mathcal{P}_U only contains

the zero-variance random variable $f = 1$. Suppose finally that ρ and Υ are linear, Δ_ρ only contains the element $z = 1$ (so $\tilde{E}_\rho = 1$) and Δ_Υ only contains $\gamma = -c \in L^2(\mathbb{P}_0)$ with $c \geq 0$ (so Υ is given by (19)). Denote $c_i = c(i) \geq 0$, $i = 1, 2, 3, 4, 5$, and take $w = 3$ in (25).

Since $\tilde{E}_\rho = 1$ and $\mathbb{P}_0(z^* + wf^* = 4) = 1$, Expression (11) implies that³

$$\left\{ \begin{array}{ll} s < 1 \Rightarrow & J^* \left(\left(\tilde{E}_\rho + w \right) c \right) = 0.6 \sum_{i=1}^5 c_i, & J^*(z^* + wf^*) = 12 \\ 1 < s < 2 \Rightarrow & J^* \left(\left(\tilde{E}_\rho + w \right) c \right) = 0.6 \sum_{i=2}^5 c_i, & J^*(z^* + wf^*) = 9.6 \\ s < 2 < 3 \Rightarrow & J^* \left(\left(\tilde{E}_\rho + w \right) c \right) = 0.6 \sum_{i=3}^5 c_i, & J^*(z^* + wf^*) = 7.2 \\ s < 3 < 4 \Rightarrow & J^* \left(\left(\tilde{E}_\rho + w \right) c \right) = 0.6 (c_4 + c_5), & J^*(z^* + wf^*) = 4.8 \\ 4 < s \Rightarrow & J^* \left(\left(\tilde{E}_\rho + w \right) c \right) = 0.6 c_5, & J^*(z^* + wf^*) = 2.4 \end{array} \right.$$

Thus, (27) leads to the bang-bang optimal solution $x^*(t)$ characterized by

$$\left\{ \begin{array}{ll} t < 1, 0.6 \sum_{i=1}^5 c_i < 12 \Rightarrow & x^* = H \\ t < 1, 0.6 \sum_{i=1}^5 c_i > 12 \Rightarrow & x^* = 1 \\ 1 < t < 2, 0.6 \sum_{i=2}^5 c_i < 9.6 \Rightarrow & x^* = H \\ 1 < t < 2, 0.6 \sum_{i=2}^5 c_i > 9.6 \Rightarrow & x^* = 1 \\ 2 < t < 3, 0.6 \sum_{i=3}^5 c_i < 7.2 \Rightarrow & x^* = H \\ 2 < t < 3, 0.6 \sum_{i=3}^5 c_i > 7.2 \Rightarrow & x^* = 1 \\ 3 < t < 4, 0.6 (c_4 + c_5) < 4.8 \Rightarrow & x^* = H \\ 3 < t < 4, 0.6 (c_4 + c_5) > 4.8 \Rightarrow & x^* = 1 \\ 4 < t, 0.6 c_5 < 2.4 \Rightarrow & x^* = H \\ 4 < t, 0.6 c_5 > 2.4 \Rightarrow & x^* = 1 \end{array} \right. \quad (37)$$

Suppose that $c_i = 6$, $i = 1, 2, 3$ and $c_i = 2.5$, $i = 4, 5$. Then, (37) leads to $x^* = 1$ if $t < 2$ and $x^* = H$ if $t > 2$ which becomes the stop-loss contract $J(1 - x^*) = (t - 2)^+$ in the classic scenario $H = 0$. Nevertheless, the opposite situation may also solve (25). In fact, if the reinsurer uses a variable loading rate $\alpha(t)$ penalizing high claims such as $c_i = 1.05$, $i = 1, 2, 3$, and $c_i = 5$, $i = 4, 5$, then (37) leads to $x^* = H$ if $t < 3$ and $x^* = 1$ if $t > 3$ which becomes $J(x^*) = (t - 3)^+$ in the classic case $H = 0$.

³notice that $J^* \left(\left(\tilde{E}_\rho + w \right) c \right) (i)$ and $J^*(z^* + wf^*) (i)$ are not relevant for $i = 1, 2, 3, 4, 5$, since the Lebesgue measure of this set vanishes and $J^* \left(\left(\tilde{E}_\rho + w \right) c \right), J^*(z^* + wf^*) \in L^2[0, M]$.

Finally, more complex bang-bang solutions are also feasible. For instance, if $c_1 = 7$, $c_2 = 1$, $c_3 = 6$, $c_4 = 1$ and $c_5 = 6$, then $x^* = H$ for $t \in (1, 2) \cup (3, 4)$ and $x^* = 1$ for $t \in (0, 1) \cup (2, 3) \cup (3, 5)$.⁴ \square

5. THE OPTIMAL REINSURANCE LINEAR PROBLEM

As indicated in Remark 5, Theorem 5 implies that for appropriate \mathcal{P}_U , ρ and Υ the dual solution $(f^*, z^*, \gamma^*, \lambda_d^*, \lambda_u^*)$ of (33) may be obtained in practice by linear programming methods. Once $(f^*, z^*, \gamma^*, \lambda_d^*, \lambda_u^*)$ is known, the linear problem (36) may lead to the optimal retention $J(x^*)$ (see the proof of Theorem 7). Nevertheless, there is second method allowing us to linearize (25), and this is the focus of this section.

Suppose that the sub-gradients of ρ and Υ are given by linear constraints (robust $CVaR$, robust absolute deviation, many cases of the robust weighted $CVaR$, etc., see Theorem 2 and Remark 3). Then, we can give a linear dual of (33) (Anderson and Nash, 1987), which is the double-dual of (25) and could lead to the optimal retention too. In other words, the optimal retention could be directly obtained by solving a linear problem. In order to shorten the exposition let us focus on the robust $CVaR$, but these ideas apply in much more frameworks.

Suppose that $\rho = RCVaR_{(\mathcal{P}_U^0, \mu)}$ and $\Upsilon(y) = \alpha \mathbb{E}_{\mathbf{P}_0}(y) + \beta RCVaR_{(\tilde{\mathcal{P}}_U^0, \tilde{\mu})}(-y)$, $\alpha, \beta \geq 0$ (see (20)), where the uncertainty level of the reinsurer $\tilde{\mathcal{P}}_U^0$ is generated by the set of priors

$$\tilde{\mathcal{P}}_U = \left\{ h \in L^2(\mathbb{P}_0); 0 \leq h \leq \tilde{R}, \mathbb{E}_{\mathbf{P}_0}(h) = 1 \right\}, \quad (38)$$

where $\tilde{R} \in L^2(\mathbb{P}_0)$ and $\tilde{R} \geq 1$. As already said, if $\tilde{R} \geq R$ then the reinsurer uncertainty will be higher than the insurer one. This is a quite natural assumption because the the reinsurer information about the involved policies may be lower, but we will not impose it. In general, we will not need any relationship between R and \tilde{R} . Similarly, we will not impose any relationship between the confidence levels μ and $\tilde{\mu}$. Furthermore, for $R = 1$ ($\tilde{R} = 1$) we include the non-ambiguous insurer (reinsurer) in the analysis, and, analogously, for $\tilde{R} = 1$, $\alpha \geq 1$ and $\beta = 0$ we include the $EVPP$ as a particular case.

Take $C = (1 + w) \alpha \mathbb{E}_{\mathbf{P}_0}(J(1))$. Proposition 1, Expression (11), Theorem 2, Ex-

⁴Notice that the solution of this example is in the line of those solutions found by Cui *et al.* (2013) for risk functions and premium principles given by distortions. The interpretation of Theorem 7 is that this type of general solution still applies in ambiguous frameworks with risk functions and premium principles incorporating both risk and ambiguity aversion, as well as imposing a lower bound H which totally eliminates the reinsurer moral hazard.

pressions (20) and (21), and Theorem 5 imply that (33) becomes

$$\left\{ \begin{array}{l} \text{Max} \quad \prec H, \lambda_d \succ - \prec 1, \lambda_u \succ + (1+w) \beta \prec 1, J^*(\gamma) \succ +C \\ \left\{ \begin{array}{l} J^*(z + wf - (1+w) \beta \gamma) - \lambda_d + \lambda_u \\ = (1+w) \alpha J^*(1) \\ \mathbb{E}_{\mathbf{P}_0}(f) = 1 \\ f \leq R \\ \mathbb{E}_{\mathbf{P}_0}(z) = 1 \\ z \leq \frac{g}{1-\mu} \end{array} \right. \left\{ \begin{array}{l} \mathbb{E}_{\mathbf{P}_0}(g) = 1 \\ g \leq R \\ \mathbb{E}_{\mathbf{P}_0}(\gamma) = 1 \\ \gamma \leq \frac{h}{1-\bar{\mu}} \\ \mathbb{E}_{\mathbf{P}_0}(h) = 1 \end{array} \right. \\ \left\{ \begin{array}{l} h \leq \tilde{R} \\ \lambda_d, \lambda_u \in L^2[0, M], f, z, g, \gamma, h \in L^2(\mathbb{P}_0) \\ \lambda_d, \lambda_u, f, z, g, \gamma, h \geq 0 \end{array} \right. \end{array} \right. \quad (39)$$

Since this problem is obviously linear, we can construct its linear dual (Luenberger, 1969, or Anderson and Nash, 1987), which will be called double-dual of (25) and is given by

$$\begin{array}{ll} \text{Min} & \left\{ \begin{array}{l} C - (1+w) \alpha \mathbb{E}_{\mathbf{P}_0}(J(\lambda_1)) + \lambda_2 + \mathbb{E}_{\mathbf{P}_0}(R\lambda_3) + \lambda_4 \\ + \lambda_6 + \mathbb{E}_{\mathbf{P}_0}(R\lambda_7) + \lambda_8 + \lambda_{10} + \mathbb{E}_{\mathbf{P}_0}(\tilde{R}\lambda_{11}) \end{array} \right. \\ s.t. & \left\{ \begin{array}{l} H - \lambda_1 \leq 0 \\ \lambda_1 - 1 \leq 0 \\ wJ(\lambda_1) - \lambda_2 - \lambda_3 \leq 0 \\ J(\lambda_1) - \lambda_4 - \lambda_5 \leq 0 \\ \frac{1}{1-\mu}\lambda_5 - \lambda_6 - \lambda_7 \leq 0 \end{array} \right\} \left\{ \begin{array}{l} (1+w) \beta J(1 - \lambda_1) - \lambda_8 - \lambda_9 \leq 0 \\ \frac{1}{1-\bar{\mu}}\lambda_9 - \lambda_{10} - \lambda_{11} \leq 0 \\ \lambda_1 \in L^2[0, M], \lambda_2, \lambda_4, \lambda_6, \lambda_8, \lambda_{10} \in \mathbb{R} \\ \lambda_3, \lambda_5, \lambda_7, \lambda_9, \lambda_{11} \in L^2(\mathbb{P}_0) \\ \lambda_3, \lambda_5, \lambda_7, \lambda_9, \lambda_{11} \geq 0 \end{array} \right. \end{array} \quad (40)$$

The complementary slackness conditions between (39) and (40) are

$$\left\{ \left\{ \begin{array}{l} \lambda_d^*(\lambda_1^* - H) = 0 \\ \lambda_u^*(1 - \lambda_1^*) = 0 \\ f^*(\lambda_2^* + \lambda_3^* - wJ(\lambda_1^*)) = 0 \\ z^*(\lambda_4^* + \lambda_5^* - J(\lambda_1^*)) = 0 \\ g^*\left(\lambda_6^* + \lambda_7^* - \frac{1}{1-\mu}\lambda_5^*\right) = 0 \\ \lambda_7^*(g^* - R) = 0 \\ \lambda_9^*\left(\gamma^* - \frac{h^*}{1-\bar{\mu}}\right) = 0 \\ \lambda_{11}^*(h^* - \tilde{R}) = 0 \end{array} \right\} \left\{ \begin{array}{l} \gamma^*(\lambda_8^* + \lambda_9^* - (1+w)\beta J(1 - \lambda_1^*)) = 0 \\ h^*\left(\lambda_{10}^* + \lambda_{11}^* - \frac{1}{1-\bar{\mu}}\lambda_9^*\right) = 0 \\ \lambda_3^*(f^* - R) = 0 \\ \lambda_5^*\left(z^* - \frac{g^*}{1-\mu}\right) = 0 \end{array} \right. \right\} \quad (41)$$

The constraints of (39) and (40), along with (41), characterize the dual and the double-dual solutions

$$\mathcal{D} = (\lambda_d^*, \lambda_u^*, f^*, z^*, g^*, \gamma^*, h^*) \quad (42)$$

and

$$\mathcal{S} = (\lambda_1^*, \lambda_2^*, \lambda_3^*, \lambda_4^*, \lambda_5^*, \lambda_6^*, \lambda_7^*, \lambda_8^*, \lambda_9^*, \lambda_{10}^*, \lambda_{11}^*) \quad (43)$$

if (40) is solvable and there is no duality gap between (39) and (40) (Anderson and Nash, 1987). Conversely, if there exist (39)-feasible and (40)-feasible solutions of (41) then (40) is solvable and there is no duality gap between (39) and (40). In general, the solvability of (40) and the absence of duality gap can not be guaranteed.

Needless to say, it is important to see that the solution of (40) leads to the optimal solution x^* of (25) and the optimal retention $J(x^*)$. Proposition 8 below shows that the absence of duality gap plays a critical role.

Proposition 8. *Suppose that the double-dual problem (40) is solvable and has no duality gap with the dual problem (39). Suppose that (43) solves the double-dual (40). Then, $x^* = \lambda_1^*$ solves the optimal reinsurance problem (25).*

Proof. Consider the solution (42) of (39), whose existence is guaranteed by Theorem 5. According to Theorem 5 we must prove that (34) holds for λ_1^* and (42). The proposition assumptions imply that (41) holds and therefore the fourth and fifth conditions in (34) hold too. Thus, it only remains to see that the first, second and third conditions in (34) hold. Since the three inequalities have a similar proof, let us show that

$$\mathbb{E}_{\mathbf{P}_0}(J(\lambda_1^*)z^*) \geq \mathbb{E}_{\mathbf{P}_0}(J(\lambda_1^*)z) \quad (44)$$

holds for every $z \in \Delta_\rho$, and let us omit the two remaining cases. The fourth equation of (41) implies that $J(\lambda_1^*)z^* = z^*(\lambda_4^* + \lambda_5^*)$. The ninth equation (41) implies that $J(\lambda_1^*)z^* = z^*\lambda_4^* + \frac{g^*}{1-\mu}\lambda_5^*$. The fifth equation leads to $J(\lambda_1^*)z^* = z^*\lambda_4^* + g^*\lambda_6^* + g^*\lambda_7^*$, and the tenth one finally leads to $J(\lambda_1^*)z^* = z^*\lambda_4^* + g^*\lambda_6^* + R\lambda_7^*$. Hence, since $\mathbb{E}_{\mathbf{P}_0}(z^*) = \mathbb{E}_{\mathbf{P}_0}(g^*) = 1$,

$$\mathbb{E}_{\mathbf{P}_0}(J(\lambda_1^*)z^*) = \lambda_4^* + \lambda_6^* + \mathbb{E}_{\mathbf{P}_0}(R\lambda_7^*). \quad (45)$$

Suppose that $z \in \Delta_\rho$. The constraints of (39) imply the existence of $g \in \mathcal{P}_U$ (see (4)) such that $z \leq \frac{g}{1-\mu}$. Hence, $J(\lambda_1^*)z \leq z(\lambda_4^* + \lambda_5^*)$ due to $z \geq 0$ and the fourth constraint of (40). $J(\lambda_1^*)z \leq z\lambda_4^* + \frac{g}{1-\mu}\lambda_5^*$ due to $z \leq \frac{g}{1-\mu}$ and $\lambda_5^* \geq 0$. $J(\lambda_1^*)z \leq z\lambda_4^* + g\lambda_6^* + g\lambda_7^*$ due to $g \geq 0$ and the fifth constraint of (40). Finally, $J(\lambda_1^*)z \leq z\lambda_4^* + g\lambda_6^* + R\lambda_7^*$ due to $g \leq R$ and $\lambda_7^* \geq 0$. Hence, $\mathbb{E}_{\mathbf{P}_0}(z) = \mathbb{E}_{\mathbf{P}_0}(g) = 1$ leads to

$$\mathbb{E}_{\mathbf{P}_0}(J(\lambda_1^*)z) \leq \lambda_4^* + \lambda_6^* + \mathbb{E}_{\mathbf{P}_0}(R\lambda_7^*). \quad (46)$$

(45) and (46) imply (44). \square

Proposition 8 is useful in practice because it makes the optimal reinsurance problem become linear, but the absence of duality gap between (39) and (40) must be

guaranteed. Theorem 10 below will show that this is the case if H , R and \tilde{R} are simple functions. Recall that simple functions are given by the general expression $\sum_{i=1}^n a_i \chi_{B_i}$, $[0, M] = B_1 \cup B_2 \cup \dots \cup B_n$ being a measurable and disjoint partition, a_1, a_2, \dots, a_n being real numbers, and χ_{B_i} being the indicator of B_i , $i = 1, 2, \dots, n$. Obviously, constant functions (such as the usual ones $H = 0$, $R = \tilde{R} = 1$) are included. Moreover, recall the the set of simple functions is dense in $L^2[0, M]$ and $L^2(\mathbb{P}_0)$, and for general H , R and \tilde{R} one can always build “good simple approximations”.

In order to prove Theorem 10 we will need some concepts and instrumental results which will be briefly summarized. If $\mathcal{B}_0 \subset \mathcal{B}$ is a sub- σ -algebra of the Borel σ -algebra \mathcal{B} of $[0, M]$, then, with obvious notations, $L^2[0, M, \mathcal{B}_0]$ is a closed subspace of $L^2[0, M]$ and $L^2(\mathbb{P}_0, \mathcal{B}_0)$ is a closed subspace of $L^2(\mathbb{P}_0)$. The orthogonal projection of $L^2(\mathbb{P}_0)$ over $L^2(\mathbb{P}_0, \mathcal{B}_0)$ is given by the conditional expectation

$$\begin{aligned} L^2(\mathbb{P}_0) &\rightarrow L^2(\mathbb{P}_0, \mathcal{B}_0) \\ y &\rightarrow \mathbb{E}_{\mathbb{P}_0}(y | \mathcal{B}_0) \end{aligned}$$

and it is known that

$$\mathbb{E}_{\mathbb{P}_0}(y_1 y_2) = \mathbb{E}_{\mathbb{P}_0}(\mathbb{E}_{\mathbb{P}_0}(y_1 | \mathcal{B}_0) y_2) \quad (47)$$

for every $y_1 \in L^2(\mathbb{P}_0)$ and every $y_2 \in L^2(\mathbb{P}_0, \mathcal{B}_0)$. Besides, if \mathcal{L} is the Lebesgue measure then $\frac{\mathcal{L}}{M}$ is a probability, the orthogonal projection

$$\begin{aligned} L^2[0, M] &\rightarrow L^2[0, M, \mathcal{B}_0] \\ y &\rightarrow \mathbb{E}_{\frac{\mathcal{L}}{M}}(y | \mathcal{B}_0) \end{aligned}$$

is also given by the by a conditional expectation, and (47) becomes

$$\mathbb{E}_{\frac{\mathcal{L}}{M}}(y_1 y_2) = \mathbb{E}_{\frac{\mathcal{L}}{M}}\left(\mathbb{E}_{\frac{\mathcal{L}}{M}}(y_1 | \mathcal{B}_0) y_2\right). \quad (48)$$

If J_0 is the composition of J and the orthogonal projections above,

$$J_0 := \mathbb{E}_{\mathbb{P}_0}(- | \mathcal{B}_0) \circ J \circ \mathbb{E}_{\frac{\mathcal{L}}{M}}(- | \mathcal{B}_0),$$

then, bearing in mind that every orthogonal projection is self-adjoint, one has that

$$\begin{aligned} J_0^* &= \left(\mathbb{E}_{\mathbb{P}_0}(- | \mathcal{B}_0) \circ J \circ \mathbb{E}_{\frac{\mathcal{L}}{M}}(- | \mathcal{B}_0) \right)^* \\ &= \mathbb{E}_{\frac{\mathcal{L}}{M}}(- | \mathcal{B}_0)^* \circ J^* \circ \mathbb{E}_{\mathbb{P}_0}(- | \mathcal{B}_0)^* \\ &= \mathbb{E}_{\frac{\mathcal{L}}{M}}(- | \mathcal{B}_0) \circ J^* \circ \mathbb{E}_{\mathbb{P}_0}(- | \mathcal{B}_0). \end{aligned}$$

In other words, if $J_0 : L^2[0, M] \rightarrow L^2(\mathbb{P}_0)$ is given by

$$J_0(y) := \mathbb{E}_{\mathbb{P}_0}\left(J\left(\mathbb{E}_{\frac{\mathcal{L}}{M}}(y | \mathcal{B}_0)\right) | \mathcal{B}_0\right) \quad (49)$$

then $J_0^* : L^2(\mathbb{P}_0) \rightarrow L^2[0, M]$ is given by

$$J_0^*(y) = \mathbb{E}_{\frac{\underline{c}}{M}}(J^*(\mathbb{E}_{\mathbb{P}_0}(y|\mathcal{B}_0))|\mathcal{B}_0). \quad (50)$$

Lemma 9. Let $(\mathcal{B}_n)_{n=1}^\infty$ be a filtration of sub- σ -algebras of \mathcal{B} .

a) Let $(u_n)_{n=1}^\infty \subset L^2(\mathbb{P}_0)$ (respectively, $(u_n)_{n=1}^\infty \subset L^2[0, M]$) be a martingale adapted to the filtration $(\mathcal{B}_n)_{n=1}^\infty$. If $(u_n)_{n=1}^\infty$ is norm bounded, then there exists $u \in L^2(\mathbb{P}_0)$ (respectively, $u \in L^2[0, M]$) such that $(u_n)_{n=1}^\infty$ converges to u both in norm and almost everywhere. Moreover, $u_n = \mathbb{E}_{\mathbb{P}_0}(u|\mathcal{B}_n)$ (respectively, $u_n = \mathbb{E}_{\frac{\underline{c}}{M}}(u|\mathcal{B}_n)$) for every $n \in \mathbb{N}$.

b) If there exists $u \in L^2(\mathbb{P}_0)$ (respectively $u \in L^2[0, M]$) such that $u_n = \mathbb{E}_{\mathbb{P}_0}(u|\mathcal{B}_n)$ (respectively, $u_n = \mathbb{E}_{\frac{\underline{c}}{M}}(u|\mathcal{B}_n)$) and $\mathcal{B} = \sigma(\cup_{n=1}^\infty \mathcal{B}_n)$ coincides with the σ -algebra generated by $\cup_{n=1}^\infty \mathcal{B}_n$, then $(u_n)_{n=1}^\infty$ converges to u both in norm and almost everywhere.

c) Suppose that $\mathcal{B} = \sigma(\cup_{n=1}^\infty \mathcal{B}_n)$. Suppose that $u \in L^2(\mathbb{P}_0)$ and consider the set $A = \{\mathbb{E}_{\mathbb{P}_0}(u|\mathcal{B}_n); n \in \mathbb{N}\} \cup \{u\}$. Then, $\left(\mathbb{E}_{\frac{\underline{c}}{M}}(-|\mathcal{B}_n) \circ J^*\right)_{n=1}^\infty$ converges to J^* uniformly in A .

Proof. a) and b) are consequences of the Doob's Martingale Convergence Theorem (Durrett, 2010), so let us prove c). Assertion b) implies that $A \subset L^2(\mathbb{P}_0)$ is a compact set. Besides

$$\left\|\mathbb{E}_{\mathbb{P}_0}(-|\mathcal{B}_n) \circ J\right\| \leq \left\|\mathbb{E}_{\mathbb{P}_0}(-|\mathcal{B}_n)\right\| \|J\| \leq \|J\|$$

implies that $\left(\mathbb{E}_{\frac{\underline{c}}{M}}(-|\mathcal{B}_n) \circ J^*\right)_{n=1}^\infty$ is bounded and therefore equi-continuous. Thus, the Ascoli's Theorem implies that $\left(\mathbb{E}_{\frac{\underline{c}}{M}}(-|\mathcal{B}_n) \circ J^*\right)_{n=1}^\infty$ converges to J^* uniformly in A if $\left(\mathbb{E}_{\frac{\underline{c}}{M}}(J^*(v)|\mathcal{B}_n)\right)_{n=1}^\infty$ converges to $J^*(v)$ for every $v \in A$, which holds due to b) (Dieudonné, 1988). \square

Theorem 10. Suppose that $\mu > 0$ and $\tilde{\mu} > 0$. If H , R and \tilde{R} are simple functions then the double-dual problem (40) is solvable and has no duality gap with the dual problem (39).

Proof. It is known that the topology of $[0, M]$ is generated by a countable basis of relative open sets (Kelly, 1975), so the Borel σ -algebra \mathcal{B} is generated by a countable sequence $(B_n)_{n=1}^\infty \subset \mathcal{B}$. Consider the σ -algebra $\mathcal{B}_n = \sigma(B_1, B_2, \dots, B_n)$ for every $n \in \mathbb{N}$. It is obvious that every \mathcal{B}_n has a finite number of elements, and the filtration $(\mathcal{B}_n)_{n=1}^\infty$ satisfies

$$\mathcal{B} = \sigma(\cup_{n=1}^\infty \mathcal{B}_n). \quad (51)$$

Since H , R and \tilde{R} are simple functions we have that they generate a σ -algebra $\sigma(H, R, \tilde{R})$ with finitely many measurable sets. Thus, by replacing \mathcal{B}_n with

$$\sigma(\mathcal{B}_n \cup \sigma(H, R, \tilde{R})),$$

we obtain a new filtration, still denoted by $(\mathcal{B}_n)_{n=1}^\infty$, which also satisfies that every σ -algebra in the filtration has finitely many elements. Moreover, (51) still holds, and H , R and \tilde{R} are \mathcal{B}_n -measurable for every $n \in \mathbb{N}$.

Consider the linear optimization problem (39- n) (respectively, (40- n)) similar to (39) (respectively, (40)) if one replaces (see (49) and (50)) J , J^* , $L^2[0, M]$ and $L^2(\mathbb{P}_0)$ with $J_n = \mathbb{E}_{\mathbb{P}_0}(-|\mathcal{B}_n) \circ J \circ \mathbb{E}_{\frac{\mathcal{L}}{M}}(-|\mathcal{B}_n)$, $J_n^* = \mathbb{E}_{\frac{\mathcal{L}}{M}}(-|\mathcal{B}_n) \circ J^* \circ \mathbb{E}_{\mathbb{P}_0}(-|\mathcal{B}_n)$, $L^2[0, M, \mathcal{B}_n]$ and $L^2(\mathbb{P}_0, \mathcal{B}_n)$. It is easy to see that (40- n) is the dual problem of (39- n). Furthermore, since both problems obviously have feasible solutions, they are bounded. Finally, since \mathcal{B}_n only has finitely many elements, (39- n) and (40- n) are finite-dimensional problems ($L^2[0, M, \mathcal{B}_n]$ and $L^2(\mathbb{P}_0, \mathcal{B}_n)$ are finite-dimensional spaces), so they are solvable and their optimal values coincide for every $n \in \mathbb{N}$. Denote by φ_n the optimal value of (39- n) and (40- n), by

$$\mathcal{D}_n = (\lambda_{d,n}^*, \lambda_{u,n}^*, f_n^*, z_n^*, g_n^*, \gamma_n^*, h_n^*) \quad (52)$$

a solution of (39- n) and by

$$\mathcal{S}_n = (\lambda_{1,n}^*, \lambda_{2,n}^*, \lambda_{3,n}^*, \lambda_{4,n}^*, \lambda_{5,n}^*, \lambda_{6,n}^*, \lambda_{7,n}^*, \lambda_{8,n}^*, \lambda_{9,n}^*, \lambda_{10,n}^*, \lambda_{11,n}^*)$$

a solution of (40- n). They satisfy the constraints of (39- n) and (40- n), which are those of (39) and (40) with the obvious modifications, and they also satisfy (41- n), which coincides with (41) under straightforward modifications too. Consider the solution \mathcal{D} (see (42)) of (39) and denote by φ the optimal value of this problem (they exist owing to Theorems 2 and 5).

(i) $\varphi_n \leq \varphi_{n+1} \leq \varphi$ for every $n \in \mathbb{N}$. Indeed, let us see that $\varphi_n \leq \varphi$. (52) is clearly (39)-feasible if $(\lambda_{d,n}^*, \lambda_{u,n}^*)$ is replaced by

$$\begin{aligned} \lambda_d &= (J^*(z^* + wf^* - (1+w)(\beta\gamma^* - \alpha)))^+ \\ \lambda_u &= (J^*(z^* + wf^* - (1+w)(\beta\gamma^* - \alpha)))^- . \end{aligned}$$

Since H is \mathcal{B}_n -measurable and $y \rightarrow \mathbb{E}_{\frac{\mathcal{L}}{M}}(y|\mathcal{B}_n)$ is increasing, (48) leads to

$$\begin{aligned} &\prec H, \lambda_d \succ = \prec H, \mathbb{E}_{\frac{\mathcal{L}}{M}}(\lambda_d|\mathcal{B}_n) \succ = \\ &\prec H, \mathbb{E}_{\frac{\mathcal{L}}{M}}((J^*(z_n^* + wf_n^* - (1+w)(\beta\gamma_n^* - \alpha)))^+|\mathcal{B}_n) \succ \\ &\geq \prec H, \left(\mathbb{E}_{\frac{\mathcal{L}}{M}}(J^*(z_n^* + wf_n^* - (1+w)(\beta\gamma_n^* - \alpha))|\mathcal{B}_n)\right)^+ \succ = \prec H, \lambda_{d,n}^* \succ . \end{aligned}$$

Similarly, $- \prec 1, \lambda_u \succ \geq - \prec 1, \lambda_{u,n}^* \succ$, and (48) implies that

$$\prec 1, J^*(\gamma^*) \succ = \prec 1, \mathbb{E}_{\frac{\underline{c}}{M}}(J^*(\gamma^*) | \mathcal{B}_n) \succ.$$

Hence, $\varphi_n \leq \prec H, \lambda_d \succ - \prec 1, \lambda_u \succ + \prec 1, J^*(\gamma^*) \succ \leq \varphi$. The inequality $\varphi_n \leq \varphi_{n+1}$ may be proved with a similar method.

(ii) $(\varphi_n)_{n=1}^\infty$ converges to φ . Indeed, consider the solution (42) of (39) and denote

$$(f_n^{**}, z_n^{**}, g_n^{**}, \gamma_n^{**}, h_n^{**}) = \mathbb{E}_{\mathbf{P}_0}((f^*, z^*, g^*, \gamma^*, h^*) | \mathcal{B}_n).$$

If

$$\begin{aligned} \lambda_{d,n}^{**} &= (J_n^*(z_n^{**} + w f_n^{**} - (1+w)(\beta \gamma_n^{**} - \alpha)))^+ \\ \lambda_{u,n}^{**} &= (J_n^*(z_n^{**} + w f_n^{**} - (1+w)(\beta \gamma_n^{**} - \alpha)))^- \end{aligned}$$

then

$$(\lambda_{d,n}^{**}, \lambda_{u,n}^{**}, f_n^{**}, z_n^{**}, g_n^{**}, \gamma_n^{**}, h_n^{**})$$

is (39– n)-feasible, so the (39– n)-objective $\tilde{\varphi}_n$ at this point satisfies $\tilde{\varphi}_n \leq \varphi_n$. Then, it is sufficient to see that $(\tilde{\varphi}_n)_{n=1}^\infty$ converges to φ . Lemma 9 implies that

$$(f_n^{**}, z_n^{**}, g_n^{**}, \gamma_n^{**}, h_n^{**})_{n=1}^\infty$$

converges to $(f^*, z^*, g^*, \gamma^*, h^*)$ and $\left(\mathbb{E}_{\frac{\underline{c}}{M}}(- | \mathcal{B}_n) \circ J^*\right)_{n=1}^\infty$ converges to J^* uniformly on $(f_n^{**}, z_n^{**}, g_n^{**}, \gamma_n^{**}, h_n^{**})_{n=1}^\infty$. Hence,

$$\begin{aligned} \lim_{n \rightarrow \infty}(\lambda_{d,n}^{**}) &= \lim_{n \rightarrow \infty}(J_n^*(z_n^{**} + w f_n^{**} - (1+w)(\beta \gamma_n^{**} - \alpha)))^+ \\ &= \lim_{n \rightarrow \infty}\left(\mathbb{E}_{\frac{\underline{c}}{M}}((J^*(z_n^{**} + w f_n^{**} - (1+w)(\beta \gamma_n^{**} - \alpha))) | \mathcal{B}_n)\right)^+ \\ &= ((J^*(z^* + w f^* - (1+w)(\beta \gamma^* - \alpha))))^+. \end{aligned}$$

Analogously,

$$\lim_{n \rightarrow \infty}(\lambda_{u,n}^{**}) = ((J^*(z^* + w f^* - (1+w)(\beta \gamma^* - \alpha))))^-.$$

Consequently, bearing in mind that $\prec 1, J_n^*(\gamma_n^{**}) \succ = \prec 1, J^*(\gamma_n^{**}) \succ$ converges to $\prec 1, J^*(\gamma^*) \succ$ (see (48)),

$$\begin{aligned} \lim_{n \rightarrow \infty}(\tilde{\varphi}_n) &= \lim_{n \rightarrow \infty}(\prec H, \lambda_{d,n}^{**} \succ - \prec 1, \lambda_{u,n}^{**} \succ + (1+w)\beta \prec 1, J_n^*(\gamma_n^{**}) \succ + C) \\ &= \prec H, ((J^*(z^* + w f^* - (1+w)(\beta \gamma^* - \alpha))))^+ \succ \\ &\quad - \prec 1, ((J^*(z^* + w f^* - (1+w)(\beta \gamma^* - \alpha))))^- \succ \\ &\quad + (1+w)\beta \prec 1, J^*(\gamma^*) \succ + C = \varphi. \end{aligned}$$

(iii) The sequence $(\lambda_{1,n}^*)_{n=1}^\infty$ is norm-bounded because the constraints of (40- n) imply that $(\lambda_{1,n}^*)_{n=1}^\infty \subset [H, 1]$.

(iv) $\lambda_{2,n}^* \geq 0$ for every $n \in \mathbb{N}$. Indeed, if $\lambda_{2,n}^* = -V < 0$ the (40- n)-constraints would lead to $0 < V \leq wJ(\lambda_{1,n}^*) - \lambda_{2,n}^* \leq \lambda_{3,n}^*$, and (41- n) would imply $f_n^* = R$. Thus, $R \geq 1$ and $\mathbb{E}_{\mathbf{P}_0}(R) = \mathbb{E}_{\mathbf{P}_0}(f_n^*) = 1$ would imply $R = 1$. If this equality does not hold then we will have a clear contradiction. If $R = 1$ holds then it is easy to see that

$$(\lambda_{1,n}^*, 0, \lambda_{3,n}^* - V, \lambda_{4,n}^*, \lambda_{5,n}^*, \lambda_{6,n}^*, \lambda_{7,n}^*, \lambda_{8,n}^*, \lambda_{9,n}^*, \lambda_{10,n}^*, \lambda_{11,n}^*) \quad (53)$$

is (40- n)-feasible and the objective function remains the same, *i.e.*, (53) still solves (40- n) and its second component is non-negative.

(v) $\lambda_{4,n}^* \geq 0$ and $\lambda_{8,n}^* \geq 0$ for every $n \in \mathbb{N}$. Indeed, both cases have a similar proof, so let us deal with $\lambda_{4,n}^*$. If $\lambda_{4,n}^* = -V < 0$ the (40- n)-constraints would lead to $0 < V \leq J(\lambda_{1,n}^*) - \lambda_{4,n}^* \leq \lambda_{5,n}^*$, and (41- n) would imply $z_n^* = \frac{g_n^*}{1-\mu}$. Then $1 = \mathbb{E}_{\mathbf{P}_0}(z_n^*) = \mathbb{E}_{\mathbf{P}_0}\left(\frac{g_n^*}{1-\mu}\right) = \frac{1}{1-\mu}$ would generate the contradiction $\mu = 0$.

(vi) $\lambda_{6,n}^* \geq 0$ and $\lambda_{10,n}^* \geq 0$ for every $n \in \mathbb{N}$. Indeed, both cases have a similar proof, so let us deal with $\lambda_{6,n}^*$. If $\lambda_{6,n}^* = -V < 0$ the (40- n)-constraints would lead to $0 < V \leq \frac{1}{1-\mu}\lambda_{5,n}^* - \lambda_{6,n}^* \leq \lambda_{7,n}^*$, and (41- n) would imply $g_n^* = R$. The rest of the proof is similar to (iv).

(vii) The sequence $(\lambda_{2,n}^*, \lambda_{4,n}^*, \lambda_{6,n}^*, \lambda_{8,n}^*, \lambda_{10,n}^*)_{n=1}^\infty$ is bounded. Indeed, it is sufficient to see that there is an upper bound. (i) shows that

$$\varphi - C + (1+w)\alpha\mathbb{E}_{\mathbf{P}_0}(J_n(\lambda_{1,n}^*)) - \sum_{i=1}^2 \mathbb{E}_{\mathbf{P}_0}(R\lambda_{2i+1,n}^*) - \mathbb{E}_{\mathbf{P}_0}(\tilde{R}\lambda_{11,n}^*) \geq \sum_{i=1}^5 \lambda_{2i,n}^*.$$

Hence, $\sum_{i=1}^2 \mathbb{E}_{\mathbf{P}_0}(R\lambda_{2i+1,n}^*) + \mathbb{E}_{\mathbf{P}_0}(\tilde{R}\lambda_{11,n}^*) \geq 0$ and $\lambda_{1,n}^* \leq 1$ lead to

$$\varphi - C + (1+w)\alpha\mathbb{E}_{\mathbf{P}_0}(J_n(1)) = \varphi - C + (1+w)\alpha\mathbb{E}_{\mathbf{P}_0}(J(1)) \geq \sum_{i=1}^5 \lambda_{2i,n}^*.$$

(viii) The sequence $(\lambda_{3,n}^*, \lambda_{5,n}^*, \lambda_{7,n}^*, \lambda_{9,n}^*, \lambda_{11,n}^*)_{n=1}^\infty$ is norm-bounded. Indeed, the restrictions and objective function of (40- n) clearly imply that

$$\begin{aligned} \lambda_{3,n}^* &= (wJ_n(\lambda_{1,n}^*) - \lambda_{2,n}^*)^+ \leq wJ_n(\lambda_{1,n}^*) \leq wJ_n(1) \\ \lambda_{5,n}^* &= (J_n(\lambda_{1,n}^*) - \lambda_{4,n}^*)^+ \leq J_n(\lambda_{1,n}^*) \leq J_n(1) \\ \lambda_{7,n}^* &= \left(\frac{1}{1-\mu}\lambda_{5,n}^* - \lambda_{6,n}^*\right)^+ \leq \frac{1}{1-\mu}\lambda_{5,n}^* \leq \frac{1}{1-\mu}J_n(1) \\ \lambda_{9,n}^* &= ((1+w)\beta J_n(1 - \lambda_{1,n}^*) - \lambda_{8,n}^*)^+ \leq (1+w)\beta J_n(1 - \lambda_{1,n}^*) \leq (1+w)\beta J_n(1) \\ \lambda_{11,n}^* &= \left(\frac{1}{1-\tilde{\mu}}\lambda_{9,n}^* - \lambda_{10,n}^*\right)^+ \leq \frac{1}{1-\tilde{\mu}}\lambda_{9,n}^* \leq \frac{1}{1-\tilde{\mu}}(1+w)\beta J_n(1) \end{aligned}$$

The sequence $(J_n(1))_{n=1}^\infty$ converges to $J(1)$ (Lemma 9b)) and therefore it is norm-bounded.

(ix) The sequence $(\mathcal{S}_n)_{n=1}^\infty$ is norm-bounded. Indeed, it immediately follows from (iii) – (viii).

(x) There exists a weakly agglomeration point \mathcal{S}^{**} of $(\mathcal{S}_n)_{n=1}^\infty$. Indeed, this is a trivial consequence of (ix) and the Alaoglu's Theorem.

(xi) The agglomeration point \mathcal{S}^{**} is (40)-feasible and the objective of this problem achieves the optimal value φ of (39) at \mathcal{S}^{**} . Indeed, first, let us see that \mathcal{S}^{**} is (40)-feasible. Constraints of (40) not involving J are obvious. There are three constraints involving J , but they can be analyzed with similar methods. Thus, let us show that $J(\lambda_1^{**}) - \lambda_4^{**} - \lambda_5^{**} \leq 0$. Fix $n_0 \in \mathbb{N}$. If $n_0 \leq n$ then

$$J_n(\lambda_{n,1}^*) - \lambda_{4,n}^* - \lambda_{5,n}^* \leq 0 \Rightarrow J_n(\lambda_{n,1}^*) \leq \lambda_{4,n}^* + \lambda_{5,n}^*.$$

Since $J_{n_0} = \mathbb{E}_{\mathbf{P}_0}(-|\mathcal{B}_{n_0}) \circ J_n$, we have $J_{n_0}(\lambda_{n,1}^*) \leq \mathbb{E}_{\mathbf{P}_0}(\lambda_{4,n}^* | \mathcal{B}_{n_0}) + \mathbb{E}_{\mathbf{P}_0}(\lambda_{5,n}^* | \mathcal{B}_{n_0})$, and therefore $J_{n_0}(\lambda_1^{**}) \leq \mathbb{E}_{\mathbf{P}_0}(\lambda_4^{**} | \mathcal{B}_{n_0}) + \mathbb{E}_{\mathbf{P}_0}(\lambda_5^{**} | \mathcal{B}_{n_0})$. Thus, Lemma 9b) leads to

$$\begin{aligned} J(\lambda_1^{**}) &= \lim_{n_0 \rightarrow \infty} J_{n_0}(\lambda_1^{**}) \leq \\ &\lim_{n_0 \rightarrow \infty} \mathbb{E}_{\mathbf{P}_0}(\lambda_4^{**} | \mathcal{B}_{n_0}) + \lim_{n_0 \rightarrow \infty} \mathbb{E}_{\mathbf{P}_0}(\lambda_5^{**} | \mathcal{B}_{n_0}) \\ &= \lambda_4^{**} + \lambda_5^{**}. \end{aligned}$$

Next, let us show that the (40)-objective equals φ at \mathcal{S}^{**} . Indeed, $\mathbb{E}_{\mathbf{P}_0}(J_n(\lambda_{n,1}^*)) = \mathbb{E}_{\mathbf{P}_0}(\mathbb{E}_{\mathbf{P}_0}(J(\lambda_{n,1}^*) | \mathcal{B}_n)) = \mathbb{E}_{\mathbf{P}_0}(J(\lambda_{n,1}^*))$ implies that

$$\begin{aligned} \varphi_n &= C - (1+w)\alpha \mathbb{E}_{\mathbf{P}_0}(J_n(\lambda_{n,1}^*)) + \lambda_{n,2}^* + \mathbb{E}_{\mathbf{P}_0}(R\lambda_{n,3}^*) + \lambda_{n,4}^* \\ &+ \lambda_{n,6}^* + \mathbb{E}_{\mathbf{P}_0}(R\lambda_{n,7}^*) + \lambda_{n,8}^* + \lambda_{n,10}^* + \mathbb{E}_{\mathbf{P}_0}(\tilde{R}\lambda_{n,11}^*) = \\ &C - (1+w)\alpha \mathbb{E}_{\mathbf{P}_0}(J(\lambda_{n,1}^*)) + \lambda_{n,2}^* + \mathbb{E}_{\mathbf{P}_0}(R\lambda_{n,3}^*) + \lambda_{n,4}^* \\ &+ \lambda_{n,6}^* + \mathbb{E}_{\mathbf{P}_0}(R\lambda_{n,7}^*) + \lambda_{n,8}^* + \lambda_{n,10}^* + \mathbb{E}_{\mathbf{P}_0}(\tilde{R}\lambda_{n,11}^*). \end{aligned}$$

The left hand side of the equality above converges to φ (see (ii)), so it must equal every agglomeration point of the right hand side. In other words,

$$\begin{aligned} \varphi &= C - (1+w)\alpha \mathbb{E}_{\mathbf{P}_0}(J(\lambda_1^{**})) + \lambda_2^{**} + \mathbb{E}_{\mathbf{P}_0}(R\lambda_3^{**}) + \lambda_4^{**} \\ &+ \lambda_6^{**} + \mathbb{E}_{\mathbf{P}_0}(R\lambda_7^{**}) + \lambda_8^{**} + \lambda_{10}^{**} + \mathbb{E}_{\mathbf{P}_0}(\tilde{R}\lambda_{11}^{**}). \end{aligned}$$

(xii) Finally, since the (40)-objective at \mathcal{S}^{**} equals the (39)-objective at \mathcal{D} , both linear optimization problems are solvable and there is no duality gap between them (Anderson and Nash, 1987). \square

6. CONCLUSIONS

This paper has dealt with the optimal reinsurance problem if both insurer and reinsurer are facing risk and uncertainty, though the classical uncertainty free case is also included. The levels of uncertainty of insurer and reinsurer do not have to be identical. As a second novelty, the decision variable is not the retained (or ceded) risk, but its sensitivity (mathematical derivative) with respect to the total claims. Thus, if one imposes strictly positive lower bounds for this variable, the reinsurer moral hazard is totally eliminated. This may be an important property because stop-loss and closely related (often optimal) contracts have been criticized. Indeed, the reinsurer would not accept these solutions in practice due to the lack of incentives of the insurer to verify claims beyond some thresholds.

Three main contributions seem to be reached. Firstly, necessary and sufficient optimality conditions are given, and they apply in all the cases contained in the general setting above. Secondly, the optimal contract is often a bang-bang solution, *i.e.*, the sensitivity between the retained risk and the total claims saturates the imposed constraints. This finding also explains why stop-loss or related contracts are often optimal if appropriate constraints are not imposed. For some special cases the optimal contract might not be a bang-bang one, but there is always a bang-bang contract as close as desired to the optimal one (the optimal reinsurance is in the closure of a set described by bang-bang solutions). Thirdly, the optimal reinsurance problem is equivalent to other linear programming problem (the double-dual problem), despite the fact that risk and uncertainty (and many pricing principles) cannot be represented by linear expressions. This may be an important finding for two reasons. On the one hand, linear problems are easy to solve in practice, since there are very efficient algorithms in both finite and infinite dimensions. On the other hand, linear problems often lead to extreme solutions, which explains why the non linear optimal reinsurance problem may be solved by a bang-bang reinsurance.

Acknowledgments. Research partially supported by *Ministerio de Economía*, Spain, Grant *ECO2012 – 39031 – C02 – 01*. The usual caveat applies.

REFERENCES

- [1] Anderson, E.J. and P. Nash, 1987. *Linear programming in infinite-dimensional spaces*. John Wiley & Sons
- [2] Arrow, K. J., 1963. Uncertainty and the welfare of medical care. *American Economic Review*, 53, 941-973.
- [3] Artzner, P., F. Delbaen, J.M. Eber and D. Heath, 1999. Coherent measures of risk. *Mathematical Finance*, 9, 203-228.

- [4] Balbás, A., B. Balbás and R. Balbás, 2013. Good deals in markets with friction. *Quantitative Finance*, 13, 827–836.
- [5] Balbás, A., B. Balbás and A. Heras, 2009. Optimal reinsurance with general risk measures. *Insurance: Mathematics and Economics*, 44, 374–384.
- [6] Balbás, A., B. Balbás and A. Heras, 2011. Stable solutions for optimal reinsurance problems involving risk measures. *European Journal of Operational Research*, 214, 796–804.
- [7] Borch, K. 1960. An attempt to determine the optimum amount of stop loss reinsurance. *Transactions of the 16th International Congress of Actuaries I*, 597–610.
- [8] Bossaerts, P., P. Ghirardato, S. Guarnaschelli and W.R. Zame, 2010. Ambiguity in asset markets: Theory and experiment. *Review of Financial Studies*, 23, 1325–1359.
- [9] Cai, J., Y Fang, Z. Li and G.E. Willmot, 2012. Optimal reciprocal reinsurance treaties under the joint survival probability and the joint profitable probability. *Journal of Risk and Insurance*, 80, 145–168.
- [10] Cai, J. and K.S. Tan, 2007. Optimal retention for a stop loss reinsurance under the *VaR* and *CTE* risk measures. *ASTIN Bulletin*, 37, 1, 93–112.
- [11] Cui, W., J. Yang and L. Wu, 2013. Optimal reinsurance minimizing the distortion risk measure under general reinsurance premium principles. *Insurance: Mathematics and Economics*, 53, 74–85.
- [12] Centeno, M.L. and O. Simoes, 2009. Optimal reinsurance. *Revista de la Real Academia de Ciencias, RACSAM*, 103, 2, 387–405.
- [13] Chi, Y. and K.S. Tan, 2013. Optimal reinsurance with general premium principles. *Insurance: Mathematics and Economics*, 52, 180–189.
- [14] Dieudonné, J., 1988. *Foundations of modern analysis*. Academic Press.
- [15] Durrett, R., 2010. *Probability: Theory and examples*. Fourth Edition. Cambridge University Press.
- [16] Gilboa, I. and D. Schmeidler, 1989. Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics*, 18, 141 – 153.

- [17] Kaluszka, M., 2005. Optimal reinsurance under convex principles of premium calculation. *Insurance: Mathematics and Economics*, 36, 375-398.
- [18] Kelly, J., 1975. *General topology*. Graduate Texts in Mathematics. Springer.
- [19] Luenberger, D.G., 1969. *Optimization by vector spaces methods*. John Wiley & Sons.
- [20] Maccheroni, F., M. Marinacci and A. Rustichini, 2006. Ambiguity aversion, robustness, and the variational representation of preferences. *Econometrica*, 74, 1447-1498.
- [21] Nakayama H., Y. Sawaragi and T. Tanino, 1985. *Theory of multiobjective optimization*. Academic Press.
- [22] Ogryczak, W. and A. Ruszczyński, 2002. Dual stochastic dominance and related mean risk models. *SIAM Journal on Optimization*, 13, 60-78.
- [23] Rockafellar, R.T., S. Uryasev and M. Zabarankin, 2006. Generalized deviations in risk analysis. *Finance & Stochastics*, 10, 51-74.
- [24] Riedel, F., 2009. Optimal stopping with multiple priors. *Econometrica*, 77, 857-908.
- [25] Rudin, W., 1973. *Functional analysis*. McGraw-Hill, Inc.
- [26] Rudin, W., 1987. *Real and complex analysis*. Third Edition. McGraw-Hill, Inc.
- [27] Zhu, S. and M. Fukushima, 2009. Worst case conditional value at risk with applications to robust portfolio management. *Operations Research*, 57, 5 1155-1168.